Clinical Case Reports: Open Access

Yumed Text
OPEN ACCESS JOURNALS

Review | Vol 5 Iss 1
ISSN: 2582-5038
https://dx.doi.org/10.46527/2582-5038.205

Assessment of Child Development by the Bayley III Scale: A Systematic Review

Orioli PA[1], Johnston C[2*], Krebs VLJ[3], Francisco RPV[4] and Carvalho WB[5]

[1]Neonatologist, Master's Student in Child Health at the Department of Pediatrics, University of São Paulo (FMUSP), Brazil

[2]Intensivist Physiotherapist, Assistant Professor of Clinical Research in Neonatology\Intensive Therapy by the Department of Pediatrics, FMUSP, Brazil

[3]Physician Neonatologist, Full Professor at the Department of Pediatrics at FMUSP, Brazil

[4]Obstetrician, Adjunct Professor, Department of Obstetrics, FMUSP, Brazil

[5]Intensivist Physician, Titular Professor in Neonatology\Intensive Care at the Department of Pediatrics, FMUSP, Brazil

[*]**Corresponding author:** Cíntia Johnston, Pt RRT MBA MsC Phd, Intensivist Physiotherapist, Assistant Professor of Clinical Research in Neonatology\Intensive Therapy by the Department of Pediatrics, FMUSP, SP-Brazil, Tel: 55 -11- 982320542; E-mail: cintia.johnston@hc.fm.usp.br

Received: February 10, 2022; **Accepted:** March 10, 2022; **Published:** March 18, 2022

Abstract

Objective: To identify the accuracy of the Bayley III Scale to assess the neuropsychomotor development (DNPM) of children, as well as to identify the degree of evidence and the level of recommendation of the selected studies.

Method: This is a systematic literature review using the keyword "Bayley III" in the Medline PubMed digital health database. The following filters were used: Clinical Trial, Controlled Clinical Trial, Meta-Analysis, Randomized Controlled Trial, Review, Scientific Integrity Review, Systematic Review. Studies published in English from January 2020 to November 2021 carried out in the neonatal and pediatric age group evaluated by the Bayley III Scale were included. The PICO method (P=patient; I=intervention; C=comparison; O=outcomes) was used to search for scientific articles in the digital health database, the degree of evidence and the level of recommendation were generated through of the GRADE rating.

Results: 118 studies were found, 12 of which met the inclusion criteria of this study. Most studies (07) were classified as C evidence and low recommendation level, 02 of them were classified as D/very low, 02 as A/high and one of them as B/moderate. The age group included ranged from 6 to 36 months and 4 studies compared the accuracy of the BSID-III with the BSID-II.

Conclusions: The Bayley III Scale has been widely applied in pediatrics to assess DNPM (cognitive, language, gross and fine motor) in the age group between one and 42 months corrected age. It has a mean sensitivity of 33.3% and specificity of 98%, with a level of evidence C and a low degree of recommendation to identify delays in DNPM in pediatrics.

Keywords: ***Bayley-III; BSID-III; DNPM; Motor development; Child development; Infant; Pediatrics; Neonatology***

Citation: Orioli PA, Johnston C, Krebs VLJ, et al. Assessment of Child Development by the Bayley III Scale: A Systematic Review. Clin Case Rep Open Access. 2022;5(1):205.

1. Introduction

The detection of possible changes in children's development is essential to predict the need to institute early intervention methods, aiming to prevent or minimize cognitive, language, motor, behavioral and social problems throughout the child's life [1-3].

Bayley scales have been applied as an objective measure to identify possible neurodevelopmental delays, both at the clinical/care level and in scientific research. The main objective of these scales is to identify children with developmental delays and provide information for planning interventions that may be necessary in this context [1-4]. They must be applied by qualified and certified health professionals [2]. If of scales that can be applied in preterm newborns (PTNB), provided that the corrected age is used to define the starting point of the analysis of its results, as well as to derive the data [4].

There are no reliable statistical data on the incidence of delay in DNPM in children. The WHO estimates that 10% of the population of any country is made up of people with some degree of disability [5]. Risk factors range from diseases, perinatal complications to adverse socioeconomic conditions. The conjuncture of biological, psychosocial and environmental factors interferes in the development of children throughout their growth. Developmental delay can be a transient, evolutionary or permanent condition, so appropriate tools and trained professionals to perform periodic assessments are necessary [6].

The known scarcity of instruments with cross-cultural validation for the assessment of neuropsychomotor development (DNPM) in children in Brazil makes the Brazilian version of the Bayley-III scale, adapted to the Portuguese language in 2012 (MADASHI) [7], considered the gold standard. for this purpose.

Thus, the main objective of this systematic literature review was to identify the accuracy of the Bayley III Scale to assess the DNPM of children, as well as to identify the degree of evidence/recommendation of the selected studies.

2. Method

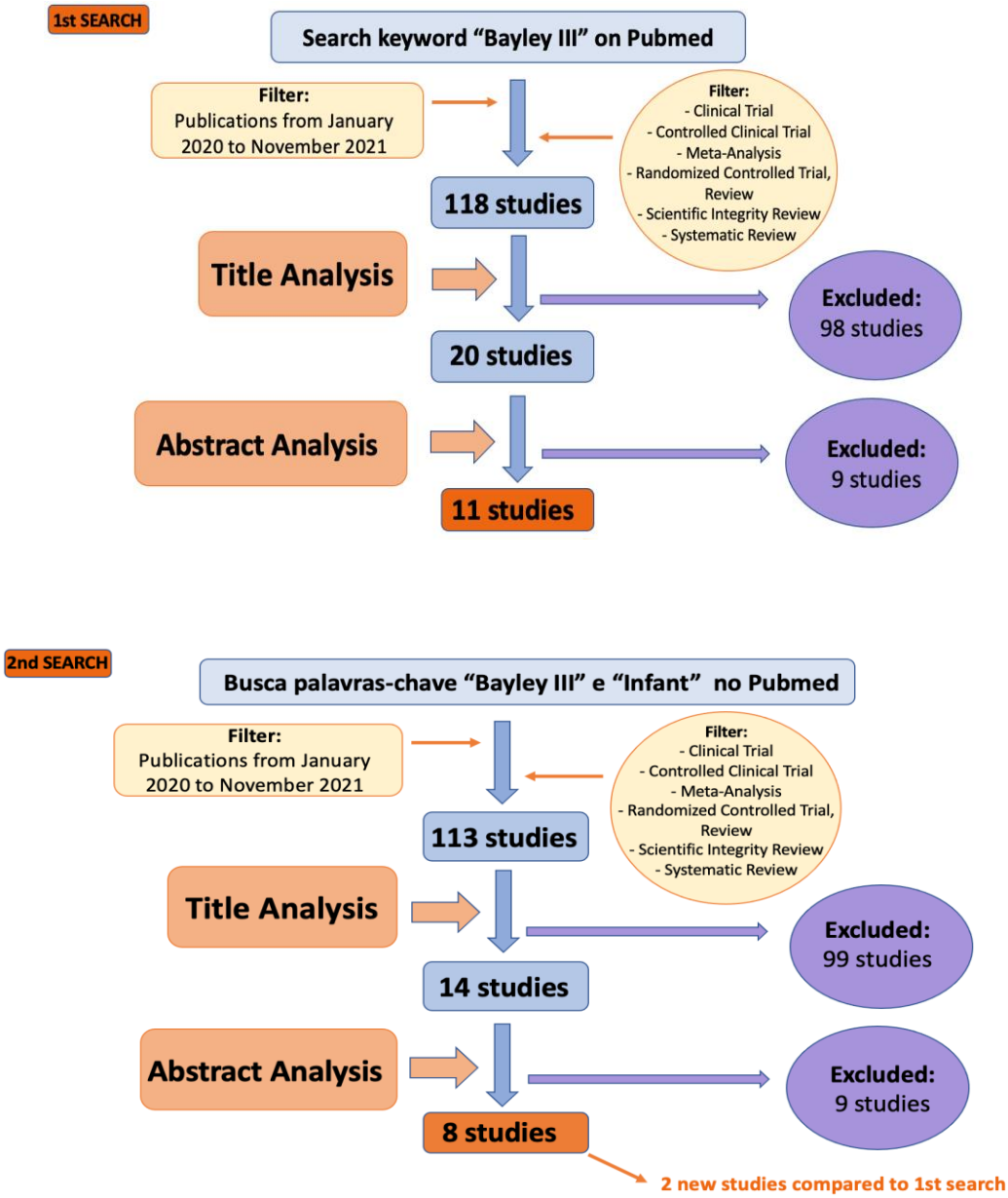
This is a systematic literature review using the keywords “Bayley III” in the Medline Pubmed digital health database. The following filters were used: Clinical Trial, Controlled Clinical Trial, Meta-Analysis, Randomized Controlled Trial, Review, Scientific Integrity Review, Systematic Review. Studies published in English from January 2020 to November 2021 carried out in the neonatal and pediatric age group evaluated by the Bayley III Scale were included.

The PICO method (P=patient; I=intervention; C=comparison; O=outcomes) was used to search for scientific articles in the digital health database [8]. The first researcher searched for scientific articles and selected by title and abstract those that met the inclusion criteria of this systematic review. The second and third researchers classified the degree of evidence/recommendation of the scientific articles found by the GRADE method [9], following the classification: A= well-designed clinical trial, B= clinical trial with mild limitation or well-designed observational study with consistent findings, C= clinical trial with moderate limitations or comparative observational study (cohort and case-control), D= clinical trial with severe limitations or observational study with limitation or non-comparative observational study. A fourth researcher analyzed the GRADE generated by the second/third researchers to verify the agreement of the degree of evidence/recommendation of

each scientific article included in this systematic review. If there was disagreement between the first two researchers, the fourth researcher (with 20 years of experience in scientific research) issued the final graduation for the study. The data are presented in a descriptive way in the form of tables and graphs.

3. Results

During the period of this systematic literature review, 118 studies were found in the digital database, through the analysis of the title, 98 studies were excluded due to not meeting the inclusion criteria (examples: studies with adults, letters to the editor, isolated use of earlier versions Bayley Scales). Thus, 20 studies remained, of which 08 were excluded after reading the abstract, due to analyzing only previous versions of Bayley III (FLOWCHART 1). The 12 studies that met the inclusion criteria are shown in TABLE 1 with their respective levels of evidence/degree of recommendation, after analysis of agreement/disagreement between the researchers involved.



4. Discussion

The first edition of the Bayley Scale of Child Development (BSID) [10] was published in 1969 and evaluated children from two to 30 months of age. The second edition, the BSID-II [11], was published in 1993 and had two assessment indices, the Mental Development Index (MDI) to assess cognitive, language and social skills; and the Psychomotor Development Index (PDI), aiming to assess fine and gross motor skills [1,2,4,12].

Although popular, the BSID and BSID-II, including the MDI, did not differentiate children with selective cognitive delay from children with language delay, just as the PDI did not differentiate children with exclusively fine motor delay from those with selectively gross delay [1,4].

Intending to address the aforementioned issues, the third edition, Bayley Scale III [13], published in 2006, includes subscales, subdivided into five assessments: Cognitive, Receptive Language, Expressive Language, Fine Motor and Gross Motor; aiming to explore more broadly the possible deviations from normality (from typical behavior) and improve the clinical utility of the Bayley scales, enabling more targeted interventions when necessary [1,4,12].

The cognitive scale is applied to assess sensory acuities and the ability to respond to them, as well as memory acquisition, learning, and problem-solving skills. The language scale assesses receptive language skills, expressive vocalizations and the beginning of verbal communication. The motor scale provides a means of assessing postural control, coordination of large muscles, and more refined hand and finger manipulation skills [13].

Unlike the two previous editions, the Bayley-III Scale, aiming at applicability in newborns and infants, was standardized using a mixed sampling procedure, that is, it includes children with typical neurological development and children with some risk factor for delayed development. development (examples: trisomy 21, cerebral palsy, prematurity, among others). With this change, the normative mean was reduced, and the Bayley-III indices were almost 7 points higher than the previous scales. This change brought into question the fact that the Bayley-III Scale underreports delays in relation to previous editions [1,2].

Over time, several studies sought to compare the performance indices of the Bayley Scales. As of 2010 [1], possible causes of the higher rates of the Bayley-III Scale began to be investigated [1]. Most studies [1,2,4,12,13,15] state (see TABLE 1) that the Bayley-III Scale overestimates the DNPM, allowing for higher scores, reports an average of 7 points higher than previous editions [4]. Consequently, studies (see TABLE 1) indicate that the Bayley-III Scale has low sensitivity (33.3%) for detecting later cognitive and motor impairment (in the medium/long term), under-identifying later cognitive and motor impairments, however, maintaining good specificity (98%) [1,3,16,17].

Several of the studies [1,2,4,12] that explored the differences between the BSID-II and the BSID-III have important methodological limitations, which limit comparisons between test versions (See comments in TABLE 1). Thus, it is not possible to conclude whether the change in scores is due to the temporal improvement in the results in the DNPM of the assessed child, since in some studies [4,12] (TABLE 1) Bayley-III was administered in its entirety, followed by the administration of additional items of the BSID-II or estimated performance of the BSID-II based on the overlap of the items. Caution is recommended when comparing results of scientific research carried out in different time periods that used different versions of these scales [2].

TABLE 1. **Studies included in the systematic review with their respective authors, type of study, level of scientific evidence, sample, interventions performed, main outcomes analyzed and main results.**

Author, Year of Publication	Type of study and Level of Scientific Evidence* (GRADE)	Sample	Interventions	Main Outcomes	Main results
Anderson PJ, et al. [1]	**Simple literature review** **D**	47 studies	Revised the psychometric properties of the Bayley-III Scale Outcomes analyzed: 1. Is the state of development overestimated? 2. Whether Bayley III is a predictor of long-term functionality?	There is considerable evidence that the Bayley-III overestimates child development, resulting in an incorrect classification of child developmental delays.	A number of strategies have been proposed to deal with inflated Bayley-III scores, none of which are considered ideal. Evidence to date suggests that Bayley-III is a poor predictor for detecting long-term cognitive and motor impairments. The Bayley-III needs new standards or alternatively it may be time for a new edition of the Bayley Scales.
Sharp MA. [2]	**Randomized crossover study** **C**	N=340 eligible =119 enrolled and randomized=77 completed the study being present in 2 assessments GI<32wks, GIc 18-22 months Weight ≤ 2000 g Eligible patients from 4 distinct clinics that are part of the Children's Hospital of Philadelphia Neonatal Care Network.	Bayley II and III were administered integrally (completely), maintaining standard of care at the clinic: - On separate visits - Random, blind, randomized order - Interval of 4 to 8 weeks between assessments - applied by different professionals	Bayley III scores were significantly higher across the range of scores and across all domains. Whether BSID-II underestimates development remains to be seen; Bayley III overestimates development or the two instruments just measure development differently.	The mean difference between the Bayley Cognitive Compound III scores and the BSID-II Mental Development Index (MDI) was nearly 1SD higher (14.1 ± 12.9 points, p <0.001). The mean difference between the Bayley Motor Compound III scores and the BSID-II Psychomotor Development Index was almost 2/3 SD (9.0 ± 11.9 p <0.001). When the severity of delay was rated using cutoff points for moderate and severe developmental delay (1 and 2 SDs below the reference norm), 40% of children (n=31/77) were rated as least severely delayed with the Bayley III (p<0.01).

Griffiths A, et al. [3]	Systematic review Search platforms: MEDLINE, Embase, CINAHL and AMED Period: between May and July 2017 **A**	N=37 studies / manuals 7 instruments met the inclusion criteria Analyzed instruments: - Bayley-III - BOT-2 - MABC-2 - SEND - NSMDA - PDMS-2 - TGMD-2	Systematically assess the psychometric properties and clinical utility of gross motor assessment tools for children aged 2 to 12 years The methodological evaluation of the documents was completed using the four-point scale of the consensus-based standards for the selection of the checklist of health status measurement instruments (COSMIN).	Identified 7 gross motor assessment tools suitable for use in clinical or research settings, each with its own strengths and limitations. Most gross motor assessments for children have validity ranging from good to excellent.	Bayley III, NSMDA and MABC-2 have evidence of predictive validity. The BOT-2, MABC-2, PDMS-2 and TGMD-2 are the most reliable assessments in this age group. Bayley III has the best predictive validity at 2 years of age for later/long term motor outcomes. None of the assessment tools demonstrates good evaluative validity. More research on evaluative gross motor assessment tools is needed.
Reuner G, et al. [4]	**Prospective cohort study** **Period: between April 2011 and March 2012** **D**	N 108 eligible babies → only 55 PT with full assessment GA All <37wks, 2/3 PT late, were evaluated between 6 months and 16 days and 7months and 15 days GAc (homogeneous group approx. 7 months GAc) Weight 43BP / 12MBP or EBP Clinical situation: PT NB treated at the perinatal center of the University Hospital of Heidelberg between May 2010 and October 2011 Infants with severe sensory impairments, cerebral palsy, genetic syndromes,	Compare Bayley-III Bayley-II, with a special focus on first-year patterns. - First, Bayley III was applied in a single session. - Raw scores for the Bayley-II cognitive and motor scales were estimated from Bayley-III items. - All examinations were performed by a trained psychologist in approximately 30 minutes.	Bayley-III scores were significantly higher than the previous edition's relevant scores in the first year of life.	- The Bayley Scales remain internationally accepted, although the third revision raises concerns. - Although all Bayley-III scores were higher than the relevant scores from the previous edition, in contrast to studies in older groups, the differences between the two test editions in the very young study group became more obvious with regarding motor performance, with the greatest difference (ten points) between the motor scales of both editions. - MBP and EBP had significantly lower motor scores than BP in both editions, - Interpretation of Bayley-III results should be based on comparison of groups rather than comparison with normative data.

		intracranial hemorrhage (ICH) > grade 1 and periventricular leukomalacia were excluded.			
Acton BV, et al. [12]	**Prospective and longitudinal cohort study** **Multicentre : 6 locations in the 4 provinces of western Canada** **C**	N 110 survivors (68% boys) Mean age 21 months (SD 4 months) + paired samples Clinical situation: - Children born June 2004 to December 2007 who survived complex cardiac surgery at 6 weeks or less - Excluded children with chromosomal alterations/needed ECMO or heart transplantation - All were followed up to 2 years of age	Report Bayley-III scores after cardiac surgery and compare with Bayley-II - Children were assessed with Bayley-III and 25 of these children completed additional items from Bayley-II	Results suggest that scores after cardiac surgery at 6 months of age or younger are generally 4 to 9 developmental quotient points below the normative values of the Bayley-III.	- The findings suggest that the outcomes for children after different heart surgeries are varied. - Bayley-III scores ranged from 1.4 (MDI / compared to language scale) to 10.0 points (MDI / compared to cognitive scale) higher than similar BSID-II scores.

Campbell SK. [13]	**Report derived from a randomized clinical trial** **C**	N 145 GA born 29-34 weeks (GImean32.4) They were evaluated at 6 without IGc (13.4 weeks of average GA) Clinical situation: No serious diseases/congenital anomalies, prenatal drug exposure Mothers with 2 environmental risk factors	Examine the concordance between the Infant Motor Performance Test (TIMP) and the Bayley III - Assessed neurodevelopment at 6 months using Bayley-III and TIMP - Babies were evaluated first with TIMP, followed by Bayley.	Despite a good correlation between TIMP and the Bayley-III Motor Component, the analysis of the results of the 2 tests in children aged 6 without IGc was widely divergent.	No children in the group were identified as having motor delay by the Bayley-III scale, while by the TIMP 41.4% were flagged for surveillance. The sensitivity of the Bayley-III motor component in agreeing with the TIMP in identifying delay is negligible at 8.3%, while the specificity is 100% (all infants scoring above the TIMP cut-off scored above the mean). at Bayley) The PPV is also 100% (5 babies scoring below average on the Bayley Motor Component had delayed TIMP scores). VPN was 61% (the vast majority of babies tested scored above average on Bayley) - Bayley's overall agreement to reflect TIMP results was 62.1%. - TIMP is preferred for the initial assessment of infants.
Duncan AF. [14]	**Secondary analysis of prospective cohort study** **Multicentric: 16 centers** **C**	N 397 PTe children – born Feb 2005 to Feb 2009 GA<28weeks	To investigate associations in PTNB (<28 wk) between neuroimaging (USG and MRI) and developmental and behavioral outcomes at 18 to 22 months - Subjected to skull USG at 3 times: 4-14days / 35-42days IGc - Subjected to MRI between 35-42 IGc - Were tested on the BIITSEA and Bayley III scales with 18-22m IGc from 2006 to 2011	Social-emotional competence contributes to deficits in cognitive and language development. The presence of MRI lesion (which includes cerebellar lesions) is associated with later socio-emotional competence and may be a useful predictor to guide evaluation and early intervention.	The presence of lesions on short-term MRI that included cerebellar lesions was significantly associated with lower BITSEA competence, but not with problem scores. Competence scores were inversely related to the presence/significance of injuries. BITSEA Problem Scale scores and positive screenings were not associated with any US or MRI findings. Positive screens on competence scores and on Competence and Problem scores were significantly associated with Bayley-III cognitive and language scores <85

			- Examined associations of problem and competency scores and positive screening rates with USG and MRI		
Martin AJ, et al. [15]	**Secondary analysis of a double-blind, randomized, controlled trial Multicentric: Canada and New Zealand** **C**	N 204 babies with suspected or confirmed NNS= 100 girls Weight median weight of 911 g median GA birth 27 weeks Clinical situation: NB with suspected or proven NNS, assessed at 24 months of age	Compare PARCA-R with BSID-III. - PARCA-R was sent to parents to be completed approximately 4 weeks before the child reaches 24 months of age of IGc. - The BSID III was administered by a certified psychologist or other trained evaluator at the time of the scheduled 24-month follow-up visit.	The results support the PARCA-R as a practical tool for the identification of appreciable cognitive and language delay at 24 months among critically ill preterm and LBW infants. Bayley-III reference norms may tend to underestimate cognitive delay.	Spearman's correlation between the PARCA-R and BSID scales was 0.43 for cognition and 0.71 for language. PARCA-R successfully predicted cases of cognitive delay and language delay with the area under the ROC curves ranging from 0.83 to 0.97 depending on the reference norms used.

Kara OK, et al. [16]	**Randomized controlled and blinded clinical trial** **Period: between August 2015 and September 2016** **B**	N 42PT=32 children (16 per group) GI 3 m GIc Weight <or=1500 g Clinical situation: - Babies with movement. general abnormal - Excluded congenital malformations They were divided into 2 groups: - Family-based intervention, consisting of a physiotherapeutic component and a family component - Traditional early intervention group	To determine the effect of family-based intervention on the motor function of preterm infants Both groups received a treatment program based on a neurodevelopmental approach during 3 to 12 months of age. - Study group: physical therapist coach, applied 2x/week for 60 min at the child's home for 9 m - Control group: 72 routine infant therapy sessions of 1h duration, 2x/week, over the 9 months The groups were evaluated on 5 occasions: at 3 m, 6 m, 9 m, 12 m and 24 m of age by a pediatric physical therapist=Applied Bayley-III	Early physical therapy can support the development of fine and gross motor function in premature babies in the first year of life (with and without atypical movements).	This study demonstrates the importance of the role of family care in the fine and gross motor development of premature babies. Neurodevelopmental improvements in fine and gross motor areas were not different when comparing the family-based group and the traditional early intervention group.
dos Santos LES, et al. [17]	**Meta-analysis Prism Design Search platforms: Cochrane Library, PubMed, PsychINFO and CINAHL** **Period: English language studies published before March 2013** **A**	N 16 studies met the inclusion criteria=1792 children GI<or=32 weeks and/or Weight<or=1500 g 8 studies investigated Bayley-I, 7 Bayley-II, 1 Bayley-III	To determine the predictive value of the mental and motor scales of editions I, II and III of the Bayley scale for the later functioning of very premature children / MBP - Children were tested using Bayley I, II or III and reassessed later using another standardized test of any developmental aspect (not the Bayley) - The predictive value of the Bayley scales was reported or calculated in	The Bayley I, II and III scales for the later development of very premature infants/MBP is limited, with the Mental Development Index (MDI) explaining 37% of later cognitive functioning and the Psychomotor Development Index explaining 12% of motor functioning. later.	Association between MDI and cognitive functioning was not significantly influenced by BW, GA, age, interval between assessments, date of birth, sex, version of the Bayley scale, PC and study quality. The predictive value of the MDI for posterior motor and language function was inconsistent across included studies. The predictive value of the Bayley Motor Scale for long-term cognitive and language functions was inconsistent across included studies. It is debatable to what extent clinical practice should be based on a single behavioral

			terms of: sensitivity, specificity, PPV/NPV, RR/OR, correlational and/or regression coefficients, relating the Bayley scales to the follow-up test. - The follow-up time of the children ranged from 3 to 14 years		assessment, such as the Bayley Scales I, II and III. Multiple longitudinal assessments are needed to describe the course of children's development and support expectations for future development.
Pogribna U, et al. [18]	**Prospective cohort study** **C**	N 50 EBP and 16 healthy RNT GI PT<1000 g, RNT, AIG Clinical situation: - Excluded with congenital CNS anomalies and those who were ventilated with unstable clinical status / term with perinatal suffering or complications or GA 42 weeks	Investigate: 1) whether regional DTI abnormalities in EBP infants (BW ≤ 1000 g) are predictive of Bayley III scores at 18 to 22 months IGc 2) to compare microstructural white matter development and neurodevelopmental outcomes of EBP infants with healthy term controls. - Were submitted to brain DTI with 38 without IGc or before discharge - Bayley's mental scale score was the main result of interest	DTI microstructural biomarkers at term-equivalent age were independent predictors of cognitive and language development at the corrected age of 18 to 22 months.	Up to 50% of EBP babies develop cognitive impairments by age 2 years.
Atukunda P, et al. [19]	**Open trial of random intervention by group** **C**	N 511 mother and child pairs IG recruited at 6-8 months, re-enrolled when children were 24 months Clinical situation: (Intervention n=77, control =78)	Information on nutrition, stimulation and hygiene education was given to mothers in the intervention group, while the control group received routine health care. Developmental results were evaluated with Bayley-III.	The maternal education intervention had positive effects on the child's development and growth at three years but did not change the composition of the gut microbiota.	Up to 50% of EBP babies develop cognitive impairments by age 2 years. The intervention group had higher Bayley-III scores than controls. An improvement in the intervention compared to the control group was obtained for the ASQ and MSEL scores. The mean difference in height-for-age z-score was greater in the intervention

			Developmental outcomes were also assessed using the Ages and Stages Questionnaire (ASQ) and the Mullen Scales of Early Learning (MSEL). Other outcomes included growth and composition of the gut microbiota.		compared to the control group: 0.50 (0.25-0.75, P = 0.0001). The composition of the gut microbiota did not differ significantly between the two study groups.

LEGEND: *GRADE: classification of the level of evidence in A/B/C/D/E, with respective degrees of recommendation (A = high; B= moderate; C= low; D= very low); N: sample number; GA: gestational age; without: weeks; GAc: gestational age corrected for age; g: grams; BSID-II: Bayley Scales of Infant Development II (BAYLEY II); MDI: Mental Development Index = Mental Development Index; SD: mean standard deviation; BOT-2: Bruininks-Oseretsky Test of Motor Proficiency 2; MABC-2: Movement Assessment Battery for Children-2; MAND: McCarron Assessment of Neuromuscular Development; NSMDA: Sensory Motor Neurodevelopmental Assessment; PDMS-2: Peabody 2 Developmental Motor Scales; TGMD-2: Gross Motor Development Test 2; COSMIN: Consensus-based standards for the selection of health status measurement instruments; EN: premature; BP: low birth weight; MBP: very low birth weight; EBP: extremely low birth weight; ICH: intracranial hemorrhage; ECMO: extracorporeal membrane oxygenation; TIMP: infant motor performance test; PPV: positive predictive value; NPV: negative predictive value; PTe: extreme preterm; USG: ultrasound; MRI: nuclear magnetic resonance; BITSEA: Brief Social Emotional Assessment for Young Children; NNS: neonatal sepsis; PARCA-R: Parents' Report on Children's Abilities - Revised; BSID-III: Bayley III; PN: birth weight; RNT: term newborn; AGA: classification in relation to weight as adequate for gestational age; CNS: central nervous system; DTI: Diffusion tensor imaging, which can predict later development of cerebral palsy; ASQ: Ages and Stages Questionnaire; MSEL: Mullen Scales of Early Learning.

Studies [1,2,4,12] comparing BSID-II with BSID-III include high-risk infants aged between six and 22 months. Thus, it is unclear whether the same effects can be found in newborns and/or younger infants. Due to difficulties related to analyzing performance at young ages, it is recommended to add other sources of information when measuring cognitive functioning throughout childhood [4].

The impact of culture and language on the outcome of assessments must be considered. The use of norms from different populations can mask the level of risk of delay in a child's DNPM. Therefore, overestimating a child's functional capacity compared to a regionally inadequate reference group can result in the child not qualifying for early intervention, which can have a negative impact on their medium/long term development [12].

The Bayley-III Scale is widely applied to assess the infant/child's early developmental status. However, due to the possibility of overestimating the DNPM [1,3,4,12], which may result in an incorrect classification of child development, studies are suggested to assess its sensitivity/specificity in specific pediatric populations.

This study has some limitations, such as: it is a literature review based on a study with samples with different methodologies and cross-cultural versions specific to the countries where the BSID-III was applied; small sample number of included studies; moderate level of evidence/grade of recommendation of included studies.

Despite the studies pointing out limitations in the application of the Bayley-III, the positive aspects of the use of the Scale stand out, such as the possibility of being applied by several professionals in the health area, as long as they are properly trained; it can be used as an instrument for monitoring the DNPM of infants over time; allows comparing gains before versus after the treatment modalities of functional alterations [2,4].

5. Conclusions

The Bayley III Scale has been widely applied in pediatrics to assess child development issues (cognitive, language, gross and fine motor) in the age group between 6 and 22 months. It has an average sensitivity of 33.3%, specificity of 98%, level of evidence C and a low degree of recommendation to identify delays in DNPM in pediatrics, and there is a need for studies with a greater number of samples to scientifically strengthen its applicability. In clinical practice, it is applied worldwide due to its positive points, especially the fact that the Scale can be applied by several health professionals as a measure to monitor the functional status of children.

REFERENCES

1. Anderson PJ, Burnett A. Assessing developmental delay in early childhood - concerns with the Bayley-III scales. Clin Neuropsychol. 2017;31(2):371-81.
2. Sharp M, DeMauro SB. Counterbalanced Comparison of the BSID-II and Bayley-III at Eighteen to Twenty-two Months Corrected Age. J Dev Behav Pediatr. 2017;38(5):322-9.
3. Griffiths A, Toovey R, Morgan PE, et al. Psychometric properties of gross motor assessment tools for children: a systematic review. BMJ Open. 2018;8(10):e021734.
4. Reuner G, Fields AC, Wittke A, et al. Comparison of the developmental tests Bayley-III and Bayley-II in 7-month-old infants born preterm. Eur J Pediatr. 2013;172(3):393-400.
5. World Health Organization. Opportunities for all: a community rehabilitation project for slums. Philippines: WHO; 1995. 54 p.
6. de FátimaDornelas L, de CastroDuarte NM, de CastroMagalhães L. Delay in neuropsychomotor development: conceptual map, definitions, uses and limitations of the term. Rev Paul de Pediatr. 2015;33(1):88-
7. Vanessa M. Translation, cross-cultural adaptation and evidence of validity of BAYLEY III scales of child development in a population in the city of Barueri, São Paulo. 2012. 85 f. Dissertation (master's in psychology) - Mackenzie Presbyterian University, São Paulo, 2012.
8. Stone PW. Popping the (PICO) question in research and evidence-based practice. Appl Nurs Res. 2002;15(3):197-8.
9. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6.
10. Bayley N. Manual for the Baylay scales of infant development. San Antonio, TX: The Psychological Corporation, USA; 1969.
11. Bayley N. Baylay scales of infant development. 2nd ed. San Antonio, TX: The Psychological Corporation, USA; 1993.
12. Acton BV, Biggs WS, Creighton DE, et al. Overestimating neurodevelopment using the Bayley-III after early complex cardiac surgery. Pediatrics. 2011;128(4):e794-800.
13. Campbell SK, Zawacki L, Rankin KM, et al. Concurrent validity of the TIMP and the Bayley III scales at 6 weeks corrected age. Pediatr Phys Ther. 2013;25(4):395-401.
14. Duncan AF, Bann CM, Dempsey A, et al. Behavioral Deficits at 18-22 Months of Age Are Associated with Early Cerebellar Injury and Cognitive and Language Performance in Children Born Extremely Preterm. J Pediatr. 2019;204:148-156.e4.

15. Martin AJ, Darlow BA, Salt A, et al. Performance of the Parent Report of Children's Abilities-Revised (PARCA-R) versus the Bayley Scales of Infant Development III. Arch Dis Child. 2013;98(12):955-8.
16. Kara OK, Sahin S, Yardimci BN, et al. The role of the family in early intervention of preterm infants with abnormal general movements. Neurosciences. 2019;24(2):101-09.
17. dos Santos ES L, de Kieviet JF, Königs M, et al. Predictive value of the Bayley scales of infant development on development of very preterm/very low birth weight children: a meta-analysis. Early Hum Dev. 2013;89(7):487-96.
18. Pogribna U, Burson K, Lasky RE, et al. Role of diffusion tensor imaging as an independent predictor of cognitive and language development in extremely low-birth-weight infants. AJNR Am J Neuroradiol. 2014;35(4):790-6.
19. Atukunda P, Muhoozi GKM, van den Broek TJ, et al. Child development, growth and microbiota: follow-up of a randomized education trial in Uganda. J Glob Health. 2019;9(1):010431.