Use of the Peabody Developmental Motor Scales – 2 and Bayley Scales of Infant Development II with Premature Infants and Toddlers

Judith C. Rhodes

A.T. Still University

Abstract

This critical analysis of relevant studies was completed to address this clinical question: In infants and toddlers born prematurely, was the Bayley Scales of Infant Development II (BSID-II) or the Peabody Developmental Motor Scales 2 (PDMS-2) more effective at detecting gross motor delays. Three databases were searched for relevant articles to answer the PICO question. In total, 587 articles were retrieved, 7 selected for consideration, and 4 were chosen. The studies revealed general poor correlation between the PDMS-2 and BSID-II with select correlation in certain areas, mostly between age-equivalent scores (AE) over standard scores. It was concluded that for the patient in this case, the tests be used in unison with clinical expertise to determine developmental delay. Further studies of sensitivity and specificity of each test was warranted.

Use of the Peabody Developmental Motor Scales – 2 and Bayley Scales of Infant Development II with Premature Infants and Toddlers

The patient was born at 24 weeks gestation via cesarean section in fetal distress weighing 0.390 Kg. He stayed 124 days in the neonatal intensive care unit (NICU). His Apgar was 6 at 5 minutes. Below 7 was an indication that the infant needed medical attention (MedlinePlus Medical Encyclopedia, 2013). During his NICU stay, he was diagnosed with failure to thrive, sickle cell trait, bronchopulmanary dysplasia, retinopathy of prematurity, respiratory distress syndrome, and asthma.

The patient lived in foster care with his paternal aunt as his guardian and two cousins. His parents were incarcerated. His pediatrician diagnosed developmental delay prior to admission. He walked independently one month prior to admission.

He was admitted in September 2013 to the Pediatric Prescribed Extended Care (PPEC) for nursing care. He was 18 months chronological age at admission. His guardian's gross motor concern was a reported high incidence of falls communicated after the therapist reached out to the caregiver. The patient had not received any physical therapy prior to admission. The patient was transported to the PPEC via bus, rarely brought by the caregiver. The patient was absent from the PPEC frequently.

At time of admission, due to an asthma exacerbation, he was prescribed pulmicort and albuterol. In 1-3% of pediatric patients, pulmicort caused hypertonia (Drugs.com, 2014). Other complaints were syncope and neck/back pain. Albuterol was found to cause shaking and trembling in the extremities and increased heart rate (Drugs.com, 2014). The stated side effects

were not demonstrated in the patient but he continued to be monitored for any signs or symptoms.

The patient had mild hypotonia and strength deficit noted (Long and Toscano, 2002). The patient had full active range of motion in all extremities and axial skeleton (Long and Toscano, 2002). He was assessed with the PDMS-2. The PDMS-2 assessed motor skill acquisition in children with delays (Long and Toscano, 2002). The gross motor section of the PDMS-2 had four sub-sections (reflex, stationary, locomotion, and object manipulation) with three used, depending on the child's age. The patient scored a 10 in stationary, 10 in locomotion, and 6 in object manipulation. A standard score between 8 and 12 was average and a standard score from 6-7 was below average (Folio and Fewell, 2000). The object manipulation section required balance skills to manipulate objects. The patient scored low in object manipulation, struggling with balance to kick a ball. The stationary section had test items for each month until 13 months, then again at 32 months until 72 months. From 13 until 32 months, a child was on target with stationary skills if they could perform at a 13 month skill level, which included the patient (Folio and Fewell, 2000). According to the Guide for Physical Therapy Practice, the child's physical therapy diagnosis was "impaired motor function and sensory integrity associated with nonprogressive disorders of the central nervous system-congenital origin" (American Physical Therapy Association (APTA), 2003). The PICO question was asked if the BSID-II, a commonly used norm referenced test, was more effective in detecting gross motor delays in premature infants and toddlers than the PDMS-2 (Long and Toscano, 2002).

Methods

Three databases were searched to include PubMed, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Physiotherapy Evidence Database (PEDro) on January 27th and February 5th and 6th, 2014. The search strategy was depicted in Figure 1. The criteria for selecting the articles was studies with subjects to include infants or toddlers diagnosed with or at risk for developmental delay, studies determining concurrent validity between BSID-II and PDMS-2 or Bayley Scales of Infant Development–III (BSID-III) and PDMS-2, and studies determining efficacy of BSID-III in detecting delay. The first database searched was PubMed. The Mesh terms used were "motor skills", "growth and development", "child development", "prematurity" and "developmental disabilities". The search revealed 26 articles. One article related to the question and, therefore, was chosen. The "related articles" option was selected for the article found in the first search and 229 articles were revealed. From that search, 4 article that compared the BSID-II and PDMS-2 with developmentally delayed children and had not been retrieved in earlier searches were chosen.

The second database utilized to search for relevant articles was CINAHL. Initially, the same search terms of "motor skills", "growth and development", "child development", "prematurity", and "developmental disabilities" were used but no articles were produced. The search terms "Bayley" and "Peabody" were used with the criteria of "abstract" and no limits. Nine articles were retrieved. Of the 9 articles, 4 were appropriate. Three of the four were repeated from a previous search and one was a new article that was selected. To ensure a thorough search, the term "Peabody" was placed alone with the limits of articles within the last 14 years, infants 2-23 months, and preschool children 2-5 years. The search found 34 articles but none were chosen. The final CINAHL search used "bayley" as the search term with the

limits of articles from 2000 to the present, all infants, and preschool children 2-5 years. In this final search, 106 articles were retrieved, 3 were determined to be relevant but only one was not repeated and chosen. Neither of the articles found in CINAHL were full text.

A third database search was performed in PEDro. The search term "motor development" was used at it was a term used for several articles already found. The first search revealed 244 articles in which none were relevant for this review. A second search utilizing the term "Peabody" was performed with 11 articles retrieved and, again, none relevant to the PICO question. A third search was performed with the term "bayley" that produced 34 articles of which none were relevant to the topic.

In all, the three searches revealed 587 articles. Of the 587 articles, 7 were retrieved for review. Of the 7 articles reviewed 4 were chosen based on the articles truly comparing the PDMS-2 and BSID-II in their ability to detect gross motor delay. The articles retrieved are summarized in Table 1 below.

Results

Article 1:

Provost, B., Heimerl, S., McClain, C., Kim, N., Lopez, B. R., & Kodituwakku, P. (2004). Concurrent validity of the Bayley scales of infant development II motor scale and the Peabody developmental motor scales - 2 in children with developmental delays. Pediatric *Physical Therapy*, 16, 149-156. Retrieved from DOI: 10.1097/01.PEP0000136005.41585.FE

The study by Provost et al. (2004) was a level 2b nonexperimental cross-sectional design study as it explored the agreement between results from two tests that had already been found to

have strong validity (Centre for Evidence-based Medicine [CEBM], 2014; Jewell, 2011). Provost et al. (2004) set out to determine if age-equivalent (AE) scores on the BSID-II and PDMS-2 correlated, if standard scores on the gross motor, fine motor, and total motor quotients on the PDMS-2 and psychomotor development index (PDI) on the BSID-II motor scale correlated, if AE and standard scores on the two tests had clinical agreement, and finally if the two tests similarly identified typically developing and delayed children. The subjects were 110 children with, or at risk for, developmental delays with parental consent referred to and who attended an evaluation at a developmental evaluation center. The concurrent validity of the AE scores on the PDMS-2 subscales and BSID-II motor scale were determined, as well as, the standard scores of the PDMS-2 gross motor quotient (GMQ), fine motor quotient (FMQ), and total motor quotient TMQ and standard score of the PDI on the BSID-II (Provost et al, 2004). The study was chosen because it analyzed the question set out in the PICO question and included premature children, although not exclusively premature infants and toddlers.

The authors used Pearson product moment correlation coefficients to determine the relationship or association between the standard scores and AE scores from the two tests and frequency of agreement to determine how comparable the results were (Provost et al., 2004; Jewell, 2011). The BSID-II Motor Scale and PDMS-2 subscale AE scores had high to very high correlation, see table 3 for strength of correlation coefficients used by the authors. The BSID-II motor scale (PDI) and PDMS-2 GMQ, FMQ and TMQ had moderate to high correlation with standard scores. The correlation coefficients had strong correlation in areas, however, frequency of agreement scores were less impressive. A disproportionate amount of children scoring as delayed on the BSID-II did not score delayed on the PDMS-2, only 10 to 23 percent of children

scoring two standard deviations or greater below the mean on the BSID-II scored comparably on the PDMS-2. Approximately half of typically developing children classified as developing typically on the PDMS-2 were delayed on the BSID-II (Provost et al., 2004).

The authors concluded that the results supported concurrent validity for AE scores on only certain portions of the tests. They determined that the BSID-II Motor scale and PDMS-2 Locomotion sub-scale AE scores had concurrent validity but the standard scores of the two tests had poor agreement and low concurrent validity. The authors concluded that therapists should have been guarded when determining eligibility for services if determined by standard scores (Provost et al., 2004).

The greatest limitation of this study was including AE scores in determining concurrent validity of the two tests. Firstly, the Pearson product moment correlation test was best used with interval or ratio data, not ordinal (Jewell, 2011). AE data was ordinal as it did not have equal intervals. A one month difference in development for a four month old was not equivalent to a one month difference in development of a four year old (Malone and Larrivee, 2007). In addition, AE scores were not a strong method of determining progress in development over time as the curve for progress flattens out (Maloney and Larrivee, 2007). Standard scores were stronger in determining average skills and deviation from the mean. Therefore, AE scores were not a strong indicator of development for analysis, nor was the proper test performed to analyze that data. The study did not have randomized selection nor were the investigators blinded to the clients being involved in a study which could have led to investigator bias (Jewell, 2011).

Article 2:

Connolly, B. H., Dalton, L., Smith, J. B., Lamberth, N. G., McCay, B., & Murphy, W. (2006).

Concurrent validity of the Bayley scales of infant development II (BSID-II) motor scale and the Peabody developmental motor scale II (PDMS-2) in 12-month-old infants. *Pediatric Physical Therapy*, 18, 190-196. Retrieved from DOI: 10.1097/01.pep0000226746.57895.57

Connolly et al. (2006) investigated the concurrent validity of the BSID-II and PDMS-2 in 15 typically developing children. The study was a level 2b study that had a nonexperimental cross-sectional design (CEBM, 2014; Jewell, 2011). The authors set out to compare concurrent validity of the BSID-II motor scale and the PDMS-2 gross motor sub-scale and fine motor sub-scale by looking specifically at AE score correlation between BSID-II motor scale and PDMS-2 GMQ and FMQ. They questioned at what level was the correlation between standard scores on the BSID-II motor scale and PDMS-2 GMQ and FMQ reliable. Finally, do standard score classifications agree between the two tests. The study was chosen because it analyzed the concurrent validity of the PDMS-2 and BSID-II. This study was included even though the children were typical because it included the exact versions of each test stated in the PICO question. The difference in subjects was considered in final analysis of this paper (Connolly et al, 2006).

Overall the correlations between the PDMS-2 and BSID-II AE and standard scores were low, the PDMS-2 FMQ and the BSID-II motor scale correlation was low, standard score correlation on the PDMS-2 GMQ and BSID-II motor scales was low, and standard score correlation between PDMS-2 TMQ and BSID-II motor scales was low. Correlations on the PDMS-2 sub-scales and the BSID-II motor scale for AE were low and a low negative correlation was found between AE scores for object manipulation and the BSID-II motor scale. A high

correlation was found between AE scores in the locomotion subtest and BSID-II motor scale (Connolly et al., 2006).

The authors concluded that there was not support for concurrent validity in this study in either AE or standard scores except for the locomotion subtest. They stated that use of a single test may alter the eligibility of children for services and supported the use of multiple sources of information. They felt that it would be beneficial to do further research with children of varying ages and increased degree of ethnicity.

Limitations of the study included the use of the Pearson correlation product moment test used with ordinal data (Jewell, 2011). The children were limited in diversity which may have presented selection bias in an under-represented sample (Jewell, 2011). Finally, the typical children averaged 12 months in age which may have increased the validity of results in AE as there wasn't variability in data intervals found at different ages (Jewell, 2011).

Article 3:

Connolly, B. H., McClune, N. O., & Gatlin, R. (2012). Concurrent validity of the Bayley-III and the Peabody developmental motor scale - 2. *Pediatric Physical Therapy*, 24, 345-352. Retrieved from DOI: 10.1097/PEP.0b013e318267c5cf

Connolly et al. (2012) performed an assessment to determine the concurrent validity of the PDMS-2 and the BSID-III in 48 children aged 29 days to 25 months 10 days with documented developmental delay or at risk for delay. The nonexperimental cross-sectional study was a level 2b exploratory study based on database data (CEBM, 2009; Jewell, 2011). The authors set out to determine AE score correlation between BSID-III motor scale and PDMS-2 GMQ and FMQ, whether the standard scores correlated between the BSID-III motor scale and

PDMS-2 GMQ and FMQ and, finally, did standard score and AE score classifications agree between the two tests (Connolly et al., 2012). The study was chosen to give further insight to the answer of the PICO question with the most updated version of the BSID-III.

The participants were split between four groups by age, below 6 months, 6 -12 months 15 days, 12 months 16 days – 18 months, and over 18 months. The authors found that AE correlations between all sub-scales of each test were low and non-significant in the under 6 months age group and composite scores in the BSID-III and PDMS-2 TMQ and FMQ showed moderate and significant correlation. In the 6 to 12 months group, scores showed moderate to high negative AE correlation and significance between fine motor of the BSID-III and grasp and visual motor subtest scores of the PDMS-2. A high and significant correlation between BSID-III and PDMS-2 TMQ and GMQ and low the correlation with FMQ was found. In the 6-12 months section, gross motor AE scores on the BSID-III and PDMS-2 showed low correlation and non-significance. AE correlations were low and non-significant in all groups in the 12 months to 18 months age group and composite scores were high and significant. Finally high correlations were found for the AE scores and composite scores on the BSID-III and PDMS-2 in all areas of the over 18 month group (Connolly et al., 2012).

The authors concluded that there was concurrent validity between the PDMS-II TMQ and the BSID-III composite scores for children in the demographic tested and that the tests have similar outcomes. They concluded that, when considering AE scores, the BSID-III should be used for children 19 months to 25 months over the PDMS-2 (Connolly et al., 2012).

The study limitations were the lack of blinding of examiners leading to possible investigator bias, the lack of a control group, the lack of randomization in participant selection,

and the possible cultural bias as a result of the children meeting demographic statistics for the state of Tennessee but not necessarily the United States (Jewell, 2011). In addition, children that did not speak English were excluded which is a large demographic in certain areas of the country. The use of the Pearson product moment correlation in analysis of the AE scores was less than ideal (Jewell, 2011).

Article 4:

Anderson, P. J., De Luca, C. R., Hutchinson, E., Roberts, G., Doyle, L. W., & Victorian Infant Collaborative Group (2010). Underestimation of developmental delay by the new Bayley-III scale. *Archives of Pediatric and Adolescent Medicine*, 164(4), 352-356.

Anderson et al. (2010) set out to determine the ability of the BSID-III to accurately identify developmental delays in the extremely premature population. The prospective cohort study, with 211 age-adjusted 2-year-old participants who were less than 28 weeks premature and 202 two-year-old controls, was a level 2b study as the authors did not explain the follow up of the children (Anderson et al., 2010; CEBM, 2014; Jewell, 2011). Other studies concluded that the BSID-III underestimated delay in typical and atypically developing children (Anderson et al., 2010; Acton et al., 2011). The study by Anderson et al. (2010) was chosen to investigate further the conclusion of the study by Connolly et al. (2012) determining concurrent validity of the BSID-III with the PDMS-2.

Psychologists administered the cognitive, language, and motor scales and used the composite scores to compare the premature children with study norms and the control group. The means for all composite and subtest scores for the premature group were significantly lower than the control. However, the means for the premature group were within the average range for

development, whereas the means of the control group were between 0.55 and 1.23 standard deviations (SD) above the normative mean. The percentage of detected delay in the premature group in cognitive, language, and motor was 13%, 21% and 16% respectively, well below the normative criteria, as were the control group percentages. The authors considered issues of recruitment and error in administration; however, they repeated procedures conducted 8 years earlier with the BSID-II that had great result. Also, the psychologists performing the assessment were experienced and completed the accredited training. The final conclusion was that the BSID-III underestimated delay in children (Anderson et al, 2010).

Overall the study was performed well with random selection of participants, a control group, and examiners blinded to the control group only. Limitations of the study included its poor repeatability and blinding of the experimental group. The authors did not describe the setting in which the participants were tested nor the time frame in which the testing took place. The experimental group examiners were not blinded, leaving opportunity for investigator bias (Jewell, 2011).

Discussion

Recognizing the need for intervention for disabled and at-risk children not yet school-aged, Part C of the Individuals with Disabilities Education Act was created to specifically address the early intervention needs of disabled children from 0 through 2 years of age. Part C funded children received the appropriate needed intervention services if their motor performance fell below a standard determined by the state providing service. (National Dissemination Center for Children with Disabilities (NICHCY), 2012). Effgen (2013) stated that the PDMS-2 and BSID-II were two norm-referenced standardized tests that measured delay in the above areas and

were commonly used in rehabilitative settings to justify the need for physical therapy and intervention services.

Even with the development of countless discriminative standardized tests to determine the degree of delay in at risk children, the determination of delay is still questionable as standardized tests did not always agree. The studies assessed to answer the PICO question of ability to detect delay between the BSID-II and PDMS-2 found poor correlation in the scores between the two tests when diagnosing for developmental delay. While some high correlation was found between AE scores on the PDMS-2 motor sub-scales and BSID-II motor scales, the standard scores on each test did not have the same high correlation. In addition, the majority of infants and toddlers found to be delayed on the BSID-II had average skills when assessed by the PDMS-2. The authors concluded no correlation or correlation only in certain areas between the two tests (Connolly et al., 2006; Provost et al., 2004). In addition, the statistical test used to determine correlation between AE scores and standard scores, although could be used with ordinal data, was best suited for interval or ratio data (Jewell, 2011).

Since the PDMS-2 and BSID-II did not correlate, then what? Provost et al. (2004) questioned the sensitivity and specificity of the two tests. Sensitivity, or the ability of a test to find individuals with a disorder and specificity, or the ability of a test to find those without a disorder, were important measures in a diagnostic test (Jewell, 2011). None of the studies performed tests to determine specificity and sensitivity to determine which test is more accurate (Anderson et al., 2010; Connolly et al., 2006; Connolly et al., 2012; Provost et al., 2004). Further studies to determine which test more accurately detects delay via tests of sensitivity and specificity were warranted.

In addition to the studies comparing the PDMS-2 and BSID-II, one study was assessed that investigated the BSID-II's ability to detect developmental delay and another the correlation between the PDMS-2 and BSID-III. The studies were chosen to give further evidence of the conclusions of the other two studies, because the studies directly comparing the BSID-II and PDMS-2 were limited, and because the newer version, BSID-III, was likely to be the standardized study of choice in the future. Determining use of one test over another was conflicted. Connolly et al. (2012) supported use of the Bayley-III over PDMS-2 for children 19 months to 25 months. The result of this study seemingly gave credence to good correlation between the PDMS-2 and BSID-III leading to the conclusion that either test would have provided proper evidence of delay. Also, the earlier studies revealed the BSID-II to find more delay in children than the PDMS-2 which may have seemed to be corrected in the BSID-III by the result of the study by Connolly et al. (2012). However, the study by Andersen et al., (2010) brought speculation to the use of the BSID-III as an assessment tool. The study looked specifically at two-year-old children, giving limited evidence to other ages, however the authors concluded the test significantly over-estimated progress in development in the studied children. Therefore, caution was taken when using the results of any one study to determine eligibility for intervention services.

Conclusion

The PICO question: Was the BSID-II more effective in detecting gross motor delays in premature infants and toddlers than the PDMS-2 was answered. Based on the evidence presented of low agreement between the BSID-II and PDMS-2 and lack of determination of one test performing better than the other, it was concluded that both tests be used in unison with

clinical expertise to determine need for physical therapy services. The standard scores were used to determine eligibility over AE scores. Finally, the BSID-2 was a better choice if the clinician found significant delay and impairments on visual observation of the patient. Further studies comparing the BSID-III and PDMS-2 were warranted, as well as, determining the sensitivity and specificity of each test.

References

Acton, B. V., Biggs, W. S., Creighton, D. E., Penner, K. A., Switzer, H. N., Petrie Thomas, J. H., . . . Robertson, C. M. (2011). Overestimating neurodevelopment using the Bayley-III after early complex cardiac surgery. *Pediatrics*,128(4), e794-e801. Retrieved from DOI: 10.1542/peds.2011-0331

American Physical Therapy Association (2003). *Guide to physical therapist practice*. Alexandria, VA: American Physical Therapy Association.

Anderson, P. J., De Luca, C. R., Hutchinson, E., Roberts, G., Doyle, L. W., & Victorian Infant Collaborative Group (2010). Underestimation of developmental delay by the new Bayley-III scale. *Archives of Pediatric and Adolescent Medicine,* 164(4), 352-356.

APGAR: MedlinePlus Medical Encyclopedia. (2013). Retrieved from http://www.nlm.nih.gov/medlineplus/ency/article/003402.htm

Connolly, B. H., Dalton, L., Smith, J. B., Lamberth, N. G., McCay, B., & Murphy, W. (2006). Concurrent validity of the Bayley scales of infant development II (BSID-II) motor scale and the Peabody developmental motor scale II (PDMS-2) in 12-month-old infants. *Pediatric Physical Therapy*, 18, 190-196. Retrieved from DOI: 10.1097/01.pep0000226746.57895.57

Connolly, B. H., McClune, N. O., & Gatlin, R. (2012). Concurrent validity of the Bayley-III and the Peabody developmental motor scale - 2. *Pediatric Physical Therapy*, 24, 345-352. Retrieved from DOI: 10.1097/PEP.0b013e318267c5cf

Drugs.com (2014). *Prescription drug information, interactions & side Effects*. Retrieved from

http://www.drugs.com

Effgen, S. K. (2013). *Meeting the Physical Therapy Needs of Children (2nd)*. Philadelphia: F.A. Davis. Retrieved February 23, 2014 from http://r2library.com.p.atsu.edu/resource/title/0803619421

Co.

Folio, M. R., & Fewell, R. R. (2000). *Peabody Developmental Motor Scales Examiner's Manual* (2nd ed.). Austin, TX: Pro-Ed.

Jewell, D. V. (2011). *Guide to evidence-based physical therapist practice.* Sudbury, MA: Jones & Bartlett Learning.

Long, T., & Toscano, K. (2002). *Handbook of Pediatric Physical Therapy* (2nd ed.). Baltimore, MD: Lippincott Williams and Wilkins.

Maloney, E.S., & Larrivee L.S. (2007). Limitations of AE scores in reporting the results of norm-referenced tests. *Contemporary Issues in Communication Science and Disorders*, 34, 86-93.

National Dissemination Center for Children with Disabilities (2012). *Early Intervention, Then and Now*. Retrieved February 19, 2014 from http://nichcy.org/babies/history

Palisano, R. J. (1986). Concurrent and predictive validities of the Bayley motor scale and the Peabody developmental motor scales. *Physical Therapy*, 66(11), 1714-1719.

Portney, L.G. & Watkins, M.P. (2000). *Foundations of Clinical Research Applications to Practice* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.

Provost, B., Heimerl, S., McClain, C., Kim, N., Lopez, B. R., & Kodituwakku, P. (2004). Concurrent validity of the Bayley scales of infant development II motor scale and the

Peabody developmental motor scales - 2 in children with developmental delays. Pediatric *Physical Therapy*, 16, 149-156. Retrieved from DOI: 10.1097/01.PEP0000136005.41585.FE

Provost, B., Crowe, T. K., & McClain, C. (2000). Concurrent validity of the Bayley scales of infant development II motor scale and Peabody developmental motor scales in two-year-old children. *Physical and Occupational Therapy in Pediatrics*, 20(1), 5-18.

Table 1: *Summary of Articles Chosen for Review*

Authors	Date	Study Design	Disposition	Rationale for Disposition
Palisano	1986	Nonexperimental, longitudinal study	Reject	The study compared concurrent validity between PDMS and BSID, not there updated versions.
Provost, Crowe, & McClain	2000	Nonexperimental, cross-sectional study	Reject	The study utilized the PDMS not the PDMS-2
Provost, Heimerl, McClain, Kim, Lopez, &Kodituwakku	2004	Nonexperimental, cross-sectional study	Accept	This research directly analyzed the concurrent validity of the BSID-II and the PDMS-2.
Connolly, Dalton, Smith, Lamberth, McCay, & Murphy	2006	Nonexperimental, cross-sectional study	Accept	The subjects were typical, not atypical developing or premature children.
Anderson, De Luca, Hutchinson, Roberts, Doyle, & the Victorian Infant Collaborative Group	2010	Prospective, cohort study	Accept	The study assessed the ability of the BSID-III to detect delay in premature children.
Acton et al.	2011	Prospective, longitudinal study	Reject	The children were at risk due to cardiac defect and surgery but not due to prematurity.
Connolly, McClune, & Gatlin	2012	Nonexperimental, cross-sectional study	Accept	The study assessed concurrent validity of the Bayley-III and PDMS-2 on children who were delayed or at risk for delays.

(Jewell, 2011)

Table 2. *Study characteristics, outcomes, and levels of evidence*

Study	Level of Evidence	Subject Characteristics	Intervention	Outcome Measures	Other Information
Connolly, McClune & Gatlin (2012)	Level 2b	48 children age ranged from 29 days to 25 months with developmental delay or at risk for delay	Assessment of developmental delay with the BSID-III to determine concurrent validity	Concurrent validity between PDMS-2 TMQ and BSID-III composite scores 0.50 to 0.89	
Anderson, De Luca, Hutchinson, Roberts, Doyle, & the Victorian Infant Collaborative Group (2010)	Level 2b	211 age-adjusted 2-year-old participants less than 28 weeks premature and 202 two-year-old controls	Assessment with the BSID-III to determine validity of the test	Not significant delay in participants with mean motor composite 100.4 with standard deviation (19.6)	The Alberta Infant Motor Scales were used as gold standard to assess for delay and inclusion
Connolly, Dalton, Smith, Lamberth, McCay & Murphy (2006)	Level 2b	15 typically developing children	Participants tested for delay using the BSID-II and PDMS-2 to determine concurrent validity	Concurrent validity was not found between PDMS-2 and BSID-III with standard score correlation scores from 0.32 to 0.94	
Provost, Heimerl, McCain, Kim,Lopez & Kodituwakku (2004)	Level 2b	110 children delayed or at risk for delay aged 3-41 months.	Assessment of development using the BSID-II and PDMS-2 to determine concurrent validity	Concurrent validity found only in certain subscale AE scores to include locomotion with 100% agreement at SD of ± 6 months	

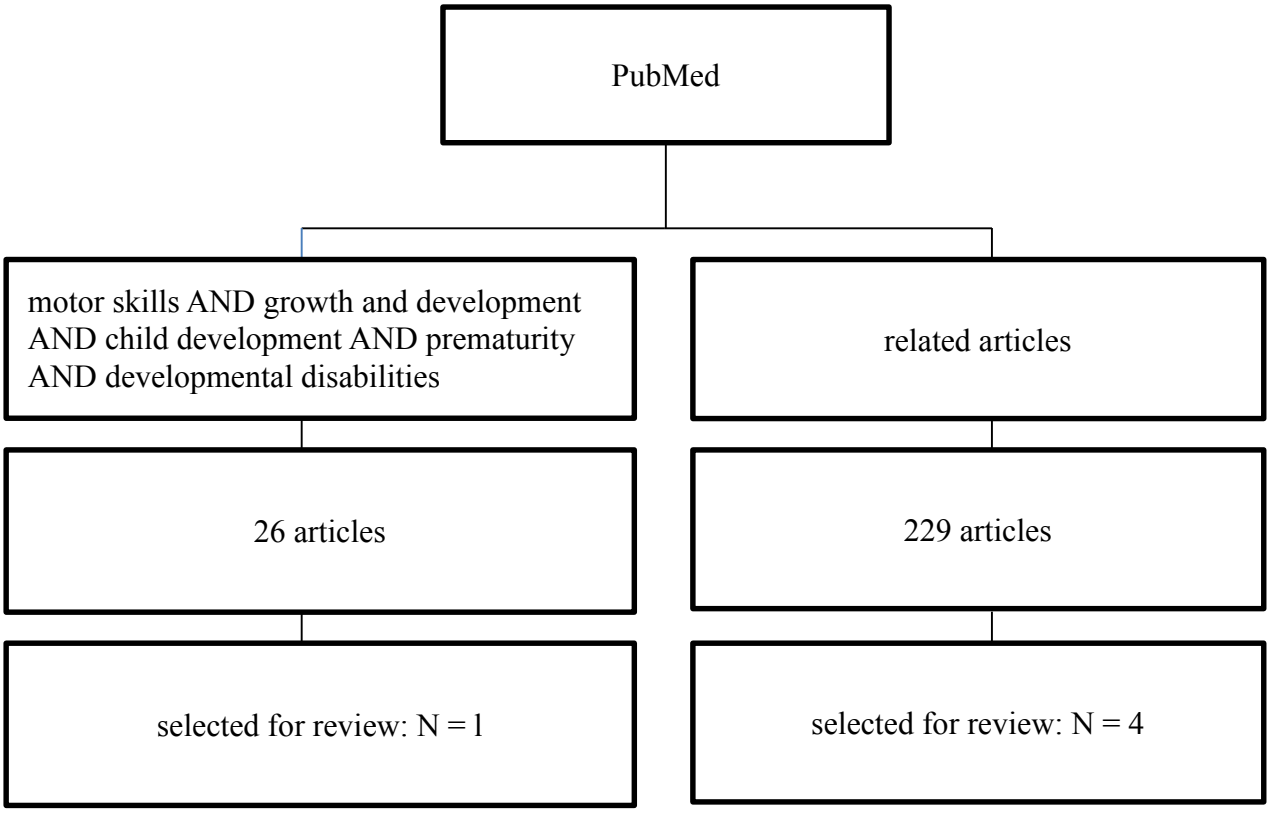
Note. BSID = Bayley Scales of Infant Development. PDMS = Peabody Developmental Motor Scales. TMQ = Total Motor Quotient. AE = Age-Equivalent. SD = Standard Deviation.

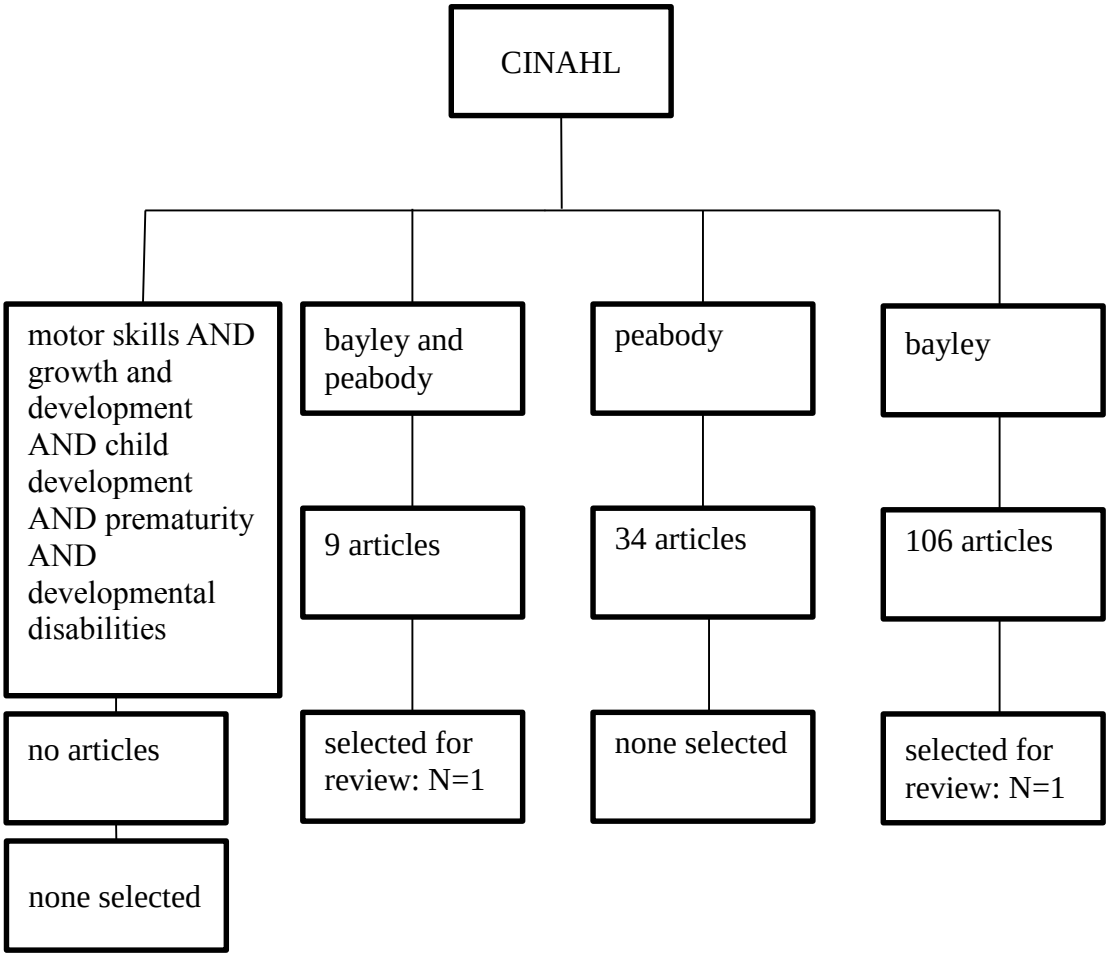
Table 3: *Strength of Correlation Coefficients*

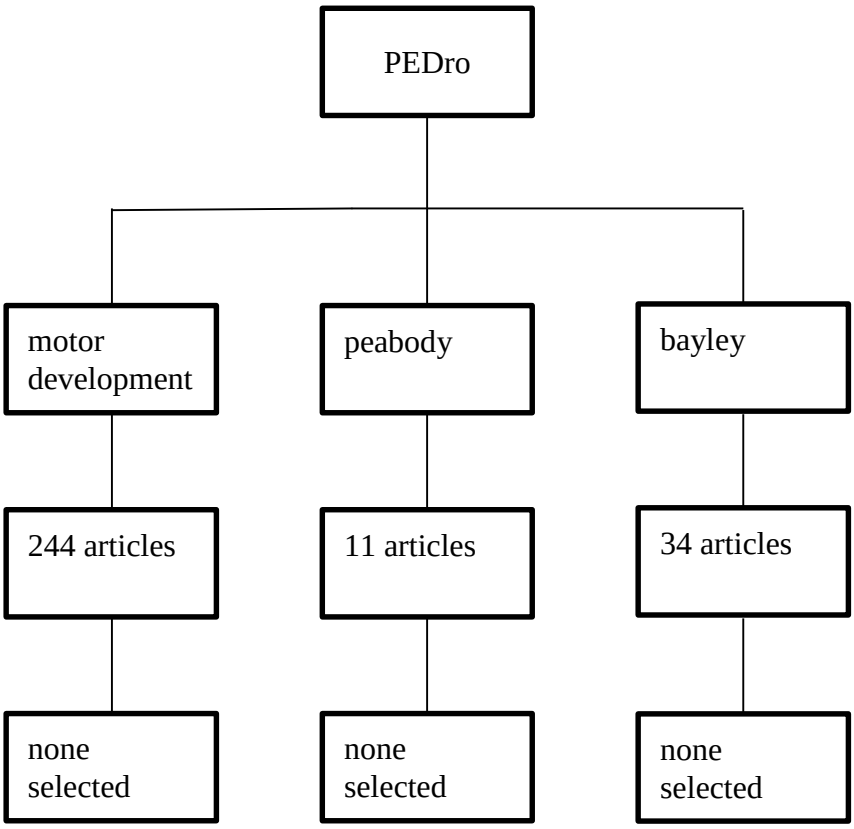
Correlation	Classification
.0 - 0.25	Little if any correlation
0.26 - 0.49	Low correlation
0.50 - 0.69	Moderate correlation
0.70 – 0.89	High correlation
0.90 – 1.0	Very high correlation

(Portney and Watkins, 2000)

Figure 1: *Summary of Database Search*




CINAHL
motor skills AND growth and development AND child development AND prematurity AND developmental disabilities
bayley and peabody
peabody
bayley
9 articles
34 articles
106 articles
no articles
selected for review: N=1
none selected
selected for review: N=1
none selected


PEDro
motor development
peabody
bayley
244 articles
11 articles
34 articles
none selected
none selected
none selected