ORIGINAL ARTICLE

Agreement between scales for screening and diagnosis of motor development at 6 months

Denise Campos,[1] Denise C. C. Santos,[2] Vanda M. G. Gonçalves,[3]
Maura M. F. Goto,[1] Amabile V. Arias,[1] Ana Carolina G. S. Brianeze,[1]
Thatiane M. Campos,[1] Bernadete B. A. Mello[1]

Abstract

Objective: To ascertain the degree of agreement between a score for screening and another for diagnosis of motor development in 6-month old infants and to define the most appropriate cutoff point for screening.

Methods: A sectional study, enrolling asymptomatic full term newborns with gestational ages from 37 to 41 weeks, who were discharged from the maternity unit 2 days after birth and are resident in the Campinas area. Infants were excluded if they presented genetic syndromes, malformations, congenital infections, intensive care admission or low birth weight. The assessment instruments investigated were the Alberta Infant Motor Scale (AIMS) and the Bayley Scales of Infant Development II (BSID-II). Two cutoff points were evaluated for the AIMS, the 5th and 10th percentiles, and for the BSID-II infants were classified according to its motor index score (IS) as having inadequate (IS < 85, at least 1 standard deviation below the mean) or adequate performance (IS ≥ 85, above the mean minus 1 standard deviation).

Results: The study sample comprised 43 infants. Six infants (14.00%) exhibited inadequate motor performance. Using the BSID-II motor classification and the 5th percentile AIMS cutoff, sensitivity was 100%, specificity 78.37%, accuracy 81.39%, kappa index 0.50 and p < 0.001; whereas, using the BSID-II motor classification and the 10th percentile AIMS cutoff, sensitivity was 100%, specificity 48.64%, accuracy 55.81%, kappa index 0.20 and p 0.025.

Conclusions: The results suggest that concordance between the two 6-month assessment scales is good. The parameters employed are best combined using the 5th percentile AIMS cutoff point.

J Pediatr (Rio J). 2006;82(6):470-4: Screening and classification, diagnosis, infant, child development, motor activity.

Introduction

Early identification of children with subtle delays and deficits can be a challenge for both clinicians and researchers, since these abnormalities manifest with the passing of time.[1] A study by Bailey et al.[2] demonstrated that families report concerns with their children's development around the seventh month of life. Diagnosis is made, on average, 1.5 months after this report and children will be referred for intervention 5 months after diagnosis.

Assessment of children's development is inefficient when based entirely on clinical impressions.[3] Less than 30% of children with mental retardation, language disorders or other developmental problems are detected based on clinical judgment.[4] Screening tests increase the rate of identification of children with suspected delays, in addition to making referral for diagnosis and intervention possible.[5,6]

In neurodevelopmental evaluation, emphasis should be placed on the use of reliable scales, with proven sensitivity and specificity, and which reflect the cultural diversity of the individuals assessed.[7] However, in Brazil, the challenge of diagnosing motor abnormalities is aggravated by the scarcity of assessment instruments standardized and validated for the local population.[8]

Screening and diagnostic scales at 6 months – Campos D et al.

Jornal de Pediatria - Vol. 82, No.6, 2006 **471**

Among the instruments utilized by Brazilian researchers, two stand out: the Bayley Scales of Infant Development II (BSID-II)[9-14] and the Alberta Infant Motor Scale (AIMS).[15-18] While they have not been validated for Brazilian children, both have been used to evaluate the motor functions both of high-risk children and those with typical development.

The BSID-II are instruments from the United States that assess children aged 1 to 42 months and that are used to diagnose development.[19] They are among the best scales available in the area of child development assessment, providing reliable, valid and precise results on the state of children's development. Their use as research instruments has received great support from the scientific community.[20]

The AIMS is a Canadian scale that evaluates gross motor function from birth up to when the ability to walk independently is acquired. It has demonstrated excellent psychometric properties with elevated rates of test-retest and interobserver reliability and validity that competes with the BSID-II. This is a screening instrument that classifies infants against a development curve between the 5th and 90th percentiles. The higher the classification percentile the lower the probability of motor development delays. Notwithstanding, interpretation of lower percentiles is less clear, suggesting greater attention to children classified at or below the 10th percentile.[21]

Screening is a rapid assessment procedure, designed to identify children who need to be referred for more detailed evaluation.[22] Screening procedures can be valuable tools since they can be applied to large populations and in situations where development is observed in a natural environment.

Based on the consideration that the AIMS does not establish an adequate range of percentiles for screening, the present study applied a screening instrument and an instrument for the diagnostic assessment of neurodevelopment to a group of healthy infants with the objective of identifying the degree of agreement between the AIMS and BSID-II during the sixth month of life and of establishing the most appropriate cutoff point for screening.

Methods

A sectional study of a cohort of infants born to full term during their sixth month of life. This project was approved by the Research Ethics Committee at the Medical Sciences Faculty (MSF) (hearing no. 087/03 – 21/10/2003) according to the provisions and principles of resolution 196/96 of the National Health Council. It was carried out in the Child Development Studies Laboratory I (CDSL-I) at the Center for Rehabilitation Studies and Research.

Neonates were enrolled from among those born live at the Neonatology Service at the Center for Integral Women's

Health Care, between May 2000 and July 2003. Newborn infants (NB) were selected if they had been born after single-fetus pregnancies, at gestational ages (GA) of 37 to 41 weeks,[23] had been in the standard maternity ward, requiring no special care with the exception of maintaining clinical and glycemic stability, were discharged from the maternity unit 2 days after birth, residents of the metropolitan area of Campinas and their parents were willing to sign the consent forms. Newborns were excluded if they were low weight (birth weight below 2,500 g), exhibited genetic syndromes, malformations or congenital infections or had been admitted to the intensive care unit.

Data to characterize the sample in terms of the variables birth weight, GA and 1st and 5th minute Apgar scores were collected during the neonatal period.

The assessment instruments employed were the motor scale from the BSID-II[19] and the AIMS.[21] The AIMS comprises 58 items illustrating the sequence of postural control development in four positions: prone (21 items), supine (9 items), sitting (12 items) and standing (16 items). During assessment, a total score was obtained by summing the items observed in each of the four positions. Each infant was then classified against a development curve varying from the 5th to the 90th percentiles. This study utilized two cutoff points: the 5th percentile or the 10th percentile.[21]

The BSID-II motor scale score was obtained from the number of tests on the 6-months routine carried out by the infant. By summing the number of tests equivalent to earlier ages, the raw score (RS) was obtained. The RS score was converted into standardized points, obtaining an index score (IS) with a mean of 100 and a standard deviation of 15.[19]

Based on their IS, infants were then classified with as having accelerated performance (IS ≥ 115), within normal performance limits (IS 85 to 114), mildly delayed performance (IS 70 to 84) or significantly delayed performance (IS ≤ 69).[19] In this study a categorical classification was adopted for motor development: adequate performance when IS ≥ 85 (greater than or equal to less than 1 standard deviation from the mean) or inadequate when IS < 85 (more than 1 standard deviation below the mean).

Infants were assessed when they had completed 6 months, plus or minus 7 days.[19] Assessments were made in the presence of their mothers or a guardian during intervals between feeds, when infants were alert and cooperative. Assessments were made simultaneously by an examiner and monitored by two observers. Responses were recorded after agreement had been reached between all three team members. The BSID-II took around 40 minutes to administer and the AIMS took 15 minutes.

Statistical analysis was performed with a 5% significance limit, using SPSS for Windows, version 11.0. The sample

was characterized by means of descriptive statistics, with neonatal variables summarized as measurements of central tendency and distribution.

Sensitivity, specificity, accuracy and kappa index were calculated for both AIMS cutoff points, using the infants' classification on the BSID-II motor scale as a gold standard. Sensitivity, specificity and accuracy were calculated using formulae adapted from Stangler et al.[24]

Results

Table 1 contains the sample characteristics at birth.

With reference to the infants' performance on the AIMS according to the two cutoff points chosen (P5 and P10), it was found that when the 5th percentile was used a majority of the infants were above the cutoff point. In contrast, when the 10th percentile was used, a majority of the infants were below the cutoff point (Table 2).

Although the majority of the infants exhibited adequate performance on the BSID-II motor scale (37 infants/ 86.00%), six infants (14.00%) were classified as having inadequate performance.

Table 3 contains the frequency of infants identified as true positives and negatives and as false positives and negatives, taking the BSID-II as gold standard. It will be observed that the two cutoff points (P5 and P10) exhibited equal proportions of false positive and negative cases. All of the infants who were classified as having inadequate motor performance by the BSID-II were

also identified as suspected cases by the AIMS. However, the 10th percentile was associated with a greater proportion of false positive cases.

Table 4 lists the degree of agreement between the AIMS and the BSID-II, for each of the cutoff points. It was confirmed that both the 5th and the 10th percentiles were sensitive for identifying at-risk infants. However, the 5th percentile exhibited greater specificity and accuracy. It was found that the degree of agreement between the AIMS and the BSID-II was greater when the 5th percentile was used as the cutoff point.

Discussion

This study exemplifies an attempt to verify an appropriate cutoff point on the AIMS for screening infants for risk of motor abnormalities. The ideal would be that the scale and method employed are able to correctly identify infants suspected of alterations and those with typical development.

Selection of a cutoff point should take into account the consequences of identifying false negative and false positive infants. False negative results delay the start of treatment. In contrast, the identification of false positive infants creates concern among parents and increases the cost of services with observation and interventions being carried out unnecessarily. Therefore, in this study, specificity was considered to be as important as sensitivity, and the AIMS cutoff point chosen as most appropriate reflects this idea.

Table 1 - Demographic characteristics of the sample at birth

Variables	n	Minimum	Maximum	Mean	Median	SD
BW (g)	43	2,500	3,850	3,065.93	3,130.00	370.26
GA (w)	42*	37.4	41.8	39.9	40.0	1.03
1 min Apgar	40**	1	10	8.18	8.50	1.72
5 min Apgar	40**	8	10	9.45	9.00	0.55

BW = birth weight; g = grams; GA = gestational age; n = number of infants; SD = standard deviation; w = weeks.
* Gestational age not available for one neonate.
† 1st and 5th minute Apgar scores unavailable for three neonates.

Table 2 - Frequency of infants classified by two AIMS cutoff points

Age	n	≤ P5 f (%)	> P5 f (%)	≤ P10 f (%)	> P10 f (%)
6 months	43	14 (32.6)	29 (67.4)	25 (58.1)	18 (41.9)

AIMS = Alberta Infant Motor Scale; f = absolute frequency; n = number of infants; % = relative frequency; P5 = 5th percentile; P10 = 10th percentile.

Screening and diagnostic scales at 6 months – Campos D et al.

Jornal de Pediatria - Vol. 82, No.6, 2006 **473**

Table 3 - Infants identified as true positives and negatives and as false positives and negatives

AIMS	True positives f (%)	False positives f (%)	False negatives f (%)	True negatives f (%)
P5	6 (14.0)	8 (18.6)	0	29 (67.4)
P10	6 (14.0)	19 (44.2)	0	18 (41.8)

AIMS = Alberta Infant Motor Scale; f = absolute frequency; % = relative frequency; P5 = 5th percentile; P10 = 10th percentile.

Table 4 - Degree of agreement between AIMS and BSID-II, by two cutoff points

Test	Sensitivity	Specificity	Accuracy	kappa index
AIMS P5	100 %	78.37 %	81.39 %	0.503*
AIMS P10	100 %	48.64 %	55.81 %	0.209[†]

AIMS = Alberta Infant Motor Scale; BSID-II = Bayley Scales of Infant Development II; P5 = 5th percentile; P10 = 10th percentile.
* p < 0.001
[†] p = 0.025

These results demonstrate that when the 10th percentile was used for AIMS assessment, a majority of the infants (58.1%) were below the cutoff point. An equally elevated proportion was found in two other studies undertaken in Brazil and using AIMS at 6 months. It was observed that 61.6% of infants born at full term with weight appropriate for GA[25] and 56% of high-risk infants[26] exhibited a percentile less than or equal to 10.

The question was therefore raised as to what factors could be responsible for the high proportion of infants that were classified at low percentiles in relation to the Canadian standard group. One hypothesis is that the rhythm of motor ability acquisition is not uniform.

This possibility is offered considering the results of authors who have compared infants from different cultural groups, demonstrating that the pattern of motor development is not universal, with evidence of characteristic behavior. Using the BSID-II, a significant difference was observed between a sample of Brazilian infants and a sample from the United States at 3, 4 and 5 months.[10]

When the manual for the Bayley scales was compiled the tests exhibited varying levels of difficulty, and for each age group those tests that 15 to 90% of children were capable of performing were selected.[19] In the age groups cited above, less than 15% of the Brazilian infants assessed performed the tests relating to rotation of the wrist and partial opposition of the thumb; sitting up momentarily, for 30 seconds and sitting up alone with stability. Concluding, the Brazilian infants in that sample exhibited low performance in tests that involved sitting and grasping abilities.[10]

Still on the subject of the non-uniform rhythm of motor acquisition, the AIMS was used with a sample of Canadian infants in order to evaluate the stability of individuals' gross motor development. In at least one of the assessments 31.1% of the infants were classified below the 10th percentile. These data suggest that a low percentile does not necessarily indicate motor delay, since development may exhibit periods of stability during which few motor abilities are acquired.[27,28]

In the current study, when a screening scale was compared to a diagnostic one, no false negative cases were detected, using the 5th or 10th percentiles. The AIMS, therefore, demonstrated high sensitivity (100%) for the detection of infants at risk of motor abnormalities.

In contrast, false positive cases were detected with both cutoff points. The 5th percentile was associated with greater specificity (78.37%) than the 10th percentile (48.64%). The specificity offered by the 5th percentile is within acceptable limits. Development screening tests

attempt to maximize both sensitivity and specificity. Generally, figures of 70 to 80% are recommended both for sensitivity and specificity, meaning that 20 to 30% of infants could be false positive cases.[29]

With respect of accuracy, a greater proportion of infants were correctly identified using the 5th percentile (81.39%) than the 10th percentile (55.81%). Similar results were found when two scales were compared (Denver-II and BSID-II), with 77.9% of children aged 3 to 36 months assessed being correctly identified.[30]

In the sample evaluated, a good level of concordance was confirmed between the AIMS and the BSID-II using the 5th percentile rather than the 10th percentile. This concordance had already been identified by Piper & Darrah,[21] evaluating 37 infants aged 4 to 7 months. A strong correlation was demonstrated between the AIMS and the BSID-II; however, the authors did not emphasize the best percentile for screening.

It can be concluded that, when the 10th percentile on the AIMS is used, the majority of infants are below the cutoff point. It is believed that the low percentile of this sample in relation to the mean of the Canadian group used as the standard can be explained by the varying rhythm of motor acquisition of these infants. There was good concordance between the evaluation instruments and the 5th percentile of the AIMS was selected as the most adequate cutoff point for screening infants for risk of motor abnormalities.

The results of this study suggest that the AIMS is a possible alternative for screening for motor development abnormalities at an early age, concentrating attention on children classified at percentiles less than or equal to 5 to 6 months of life.

References

1. Bailey DB, Skinner D, Hatton D, Roberts J. Family experiences and factors associated with the diagnosis of fragile X syndrome. J Dev Behav Pediatr. 2000;2:315-21.
2. Bailey DB, Hebbeler K, Scarborough A, Spiker D, Mallik S. First experiences with early intervention: a national perspective. Pediatrics. 2004;113:887-96.
3. Dworkin PH. Developmental screening: still expecting the impossible? Pediatrics. 1992;89:1253-5.
4. Glascoe FP. Early detection of developmental and behavioral problems. Pediatr Rev. 2000;21:272-9.
5. 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.
6. Rydz D. Developmental screening. J Child Neurol. 2005;20:4-21.
7. Santos DCC, Ravanini SG. Aspectos do diagnóstico do desenvolvimento motor. In: Moura-Ribeiro MV, Gonçalves VM. Neurologia do desenvolvimento da criança. Rio de Janeiro: Revinter; 2006. p. 258-69.
8. Burns YR, Higgins C. Roteiro do exame em fisioterapia. In: Burns YR, Macdonald J. Fisioterapia e crescimento na infância. São Paulo: Santos Editora; 1999. p. 91-111.
9. Grantham-McGregor SM, Lira PIC, Ashworth A, Morris SS, Assunção AMS. The development of low birth weight term infants and the effects of the environment in northeast Brazil. J Pediatr. 1998;132:661-6.
10. Santos DC, Gabbard C, Gonçalves VM. Motor development during the first year: a comparative study. J Genet Psychol. 2001;162:143-53.
11. Eickmann SH, Lima AC, Guerra MQ, Lima MC, Lira PI, Huttly SR, et al. Improved cognitive and motor development in a community-based intervention of psychosocial stimulation in northeast Brazil. Dev Med Child Neurol. 2003;45:536-41.
12. Gagliardo HG, Gonçalves VMG, Lima MC, Françoso MF, Aranha Neto A. Visual function and fine-motor control in small-for-gestational age infants. Arq Neuropsiquiatr. 2004;62:955-62.
13. Lima MC, Eickmann SH, Lima AC, Guerra MQ, Lira PI, Huttly SR, et al. Determinants of mental and motor development at 12 months in a low income population: a cohort study in northeast Brazil. Acta Paediatr. 2004;93:969-75.
14. Goto MM, Gonçalves VM, Netto AA, Morcillo AM, Moura-Ribeiro MV. Neurodesenvolvimento de lactentes nascidos a termo pequenos para a idade gestacional no 2º mês de vida. Arq Neuropsiquiatr. 2005;63:75-82.
15. Mancini MC, Teixeira S, Araújo LG, Paixão ML, Magalhães LC, Coelho ZA, et al. Estudo do desenvolvimento da função motora aos 8 e 12 meses de idade em crianças nascidas pré-termo e a termo. Arq Neuropsiquiatr. 2002;60:974-80.
16. Mello EQ. Aplicabilidade da escala Alberta em lactentes de risco social [dissertação]. São Paulo (SP): Universidade de São Paulo; 2003.
17. Formiga CK, Pedrazzani ES, Tudella E. Desenvolvimento motor de lactentes pré-termo participantes de um programa de intervenção fisioterapêutica precoce. Rev Bras Fisioter. 2004;8:239-45.
18. Silva PL, Santos DC, Gonçalves VM. Influência de práticas maternas no desenvolvimento motor do 6º ao 12º meses de vida. Rev Bras Fisioter. 2006;10:227-33.
19. Bayley N. Bayley scales of infant development. II Manual. San Antonio: Harcourt Brace; 1993.
20. Gabbard C, Rodrigues LP. Testes contemporâneos de avaliação do comportamento motor infantil. In: Moura-Ribeiro MV, Gonçalves VM. Neurologia do desenvolvimento da criança. Rio de Janeiro: Revinter; 2006. p. 243-57.
21. Piper MC, Darrah JM. Motor assessment of the developing infant. Philadelphia: WB Saunders; 1994.
22. American Academy of Pediatrics, Committee on Children with Disabilities. Developmental surveillance and screening of infants and young children. Pediatrics. 2001;108:192-6.
23. Organização Mundial da Saúde, CID-10. Classificação estatística internacional de doenças e problemas relacionados à saúde. São Paulo: EDUSP; 1999.
24. Stangler SR, Huber CJ, Routh DK. Screening growth and development of preschool children: a guide for test selection. New York: MacGraw-Hill; 1980.
25. Campos D. Controle postural de lactentes nascidos a termo pequenos para a idade gestacional [dissertação]. Piracicaba (SP): Universidade Metodista de Piracicaba; 2005.
26. Frônio JS. Desenvolvimento neuropsicomotor nos primeiros 18 meses de vida de lactentes de alto risco [tese]. Campinas (SP): Universidade Estadual de Campinas; 2005.
27. Darrah J, Redfern L, Maguire TO, Beaulne AP, Watt J. Intra-individual stability of rate of gross motor development in full-term infants. Early Hum Dev. 1998;52:169-79.
28. Darrah J, Hodge M, Magill-Evans J, Kembhavi G. Stability of serial assessment of motor and comunication abilities in typically developing infants - implications for screening. Early Hum Dev. 2003;72:97-110.
29. Glascoe FP. Parent's concerns about children's development: prescreening technique or screening test. Pediatrics. 1997;99:522-8.
30. Frankenburg WK, Dodds J, Archer P, Bresnick B, Maschka P, Edelman N, et al. Denver II screening manual. Denver: Denver Developmental Materials; 1990.

Correspondence:
Vanda Maria Gimenes Gonçalves
Departamento de Neurologia FCM/UNICAMP
Cidade Universitária Zeferino Vaz, Cx. Postal 6111
CEP 13081-970 – Campinas, SP – Brazil
Tel.: +55 (19) 3788.7372 – Fax: +55 (19) 3788.7483
E-mail: vandagg@uol.com.br