Dorota Mrożek-Budzyn, Małgorzata Augustyniak Agnieszka Kiełtyka, Renata Majewska

VALIDITY AND CLINICAL UTILITY OF CHILDREN DEVELOPMENT ASSESSEMENT USING MILESTONES REPORTED BY MOTHERS

Department of Epidemiology and Preventive Medicine
Jagiellonian University - Medical College

ABSTRACT

INTRODUCTION. The monitoring of infants development during preventive care visits to identify children whose development is concerning for delay is an essential part of pediatric practice.
STUDY OBJECTIVE. The aim of the study was to examine the validity and clinical utility of developmental milestones reported by mothers in assessment of children development compared with the outcomes of BSID-II (Bayley Scales of Infant Development – second edition).
MATERIAL AND METHODS. The cohort recruited prenatally, included 384 children. The Mental and Motor Scales of BSID-II were administered to each child at the end of the 12th, 24th and 36th month of life. When children were 3 years old, mothers were questioned about their child's age at attainment of 8 significant developmental milestones.
RESULTS. Sensitivity for the developmental milestones compared with score on the motor and mental scales of the BSID-II varied from 25.0% to 75.0%, specificity from 54.1% to 80.2%. The all of analysed milestones were characterized by low positive predictive value and rather high the negative one.
CONCLUSION. Parent report developmental milestones are a better tool for excluding those children who attain milestones rapidly, as a group with low risk of developmental delays, than in identifying children whose development is suspected of being delayed.

Key words: *children, development assessment, milestones, Bayley Scales*

INTRODUCTION

The monitoring of infants development during preventive care visits to identify children whose development is atypical or concerning for delay is an essential part of pediatric practice. Early identification of children with developmental delays or at risk of delay allows for referral to relevant intervention services, which have been shown to improve developmental and behavioral outcomes (1). Comparison of a child's current skills to developmental milestones data remains the most frequently reported method of development surveillance for physicians in practice, in conjunction with the physical examination of the child (2). It is therefore crucial to understand the validity, utility and limitations of developmental milestones as a tool in development surveillance. Full developmental assessment using the tests such as the Bayley Scales of Infant Development (BSID-II) are the "gold standard" but they are difficult to use on a large scale because they are expensive, time consuming and require trained staff (3). This developmental test is the widely used standardized measure of the development of infants and toddlers from 1 to 42 months of age in both clinical settings and research. This test was standardized in the U.S. and its high reliability and validity have been established. Though criticisms have been raised (4) concerning a number of methodological problems with BSID-II, the strengths of this test still outweigh its weaknesses, making it the best and most used method to assess development so far, and a useful reference tool (5).

The aim of the study was to examine the validity and clinical utility of developmental milestones reported by mothers in assessment of children development compared with the outcomes of BSID-II.

MATERIAL AND METHODS

All of the children involved in this study were part of a larger cohort study being followed in a collaborative study with Columbia University in New York on the vulnerability of fetus and child to environmental factors. The material and methods of the cohort study were published earlier (6).

When children were 3 years old, mothers were questioned about their child's age at attainment of 8 significant developmental milestones skills in the following order: lifting head while prone, sitting without support, standing without assistance, walking alone, walking upstairs, bladder trained, bowel trained and first meaningful words. The direct interview was conducted by pediatricians.

The Bayley Scales of Infants Development, second edition (BSID-II), was administered in 12th, 24th and 36th month of life (within 4 weeks of the target age). The Psychomotor Scale assesses control of gross and fine muscle groups (rolling, crawling, creeping, sitting, standing, walking, running, and jumping). The Mental Scale includes items that assess memory, habituation, problem solving, early number concepts, generalization, classification, vocalization, language, and social skills (5). Test scores are adjusted to the age of the child to obtain the Psychomotor Development Index (PDI) and the Mental Development Index (MDI). Test results are in one of four categories: 1) accelerated performance (score ≥ 115), 2) within normal limits (score, 85 to 114), 3) mildly delayed performance (score, 70 to 84), and 4) significantly delayed (score ≤ 69).

The BSID-II were conducted at the Department of Epidemiology and Preventive Medicine by trained examiners being unaware of the data about the infants age of attainment the developmental milestones.

Statistical methods. Diagnostic use of parent report milestones was determined by constructing the Receiver Operating Characteristic (ROC) curve. The motor and mental scale of BSID-II served as a reference tool of children's development. The criterion for qualifying children as one of the developmental delayed group was a standard score on the motor and mental scale of BSID-II less than 85 points. The milestones separately were used as independent variables to determine the cut-off criteria for discriminating between children with developmental delays and these functioning within or above the normal range. Area under ROC curve, the sensitivity, specificity and positive and negative predictive values were calculated for the each milestone.

RESULTS

Characteristics of study population, the means BSID-II outcomes and the age of developmental milestones attainment by children were presented in previous paper (7). The categories of development based on BSID-II were shown in table I.

The most efficient cut-off points were evaluated respectively for each of milestone: for lifting head later than in the 2nd month, sitting up later than in the 6th month, standing later than in the 9th month, walking unassisted later than in the 12th month, walking upstairs later than in the 18th month, bladder control later than in the 33rd month, bowel control later than in the 29th month and for first words being spoken later than in the 18th month.

For score on the motor scale, the area under the ROC curve was statistically higher than 0.5 in 12th month for all milestones, in 24th month for all except for sitting without assistance, in the 36th month for all except for standing without assistance. The area under the ROC curve for the score on the mental scale was significant at the all age levels for first words spoken, toilet training and walking alone. Walking upstairs produced

Table I. BSID-II scores in gender groups

Bayley Scale		Bayley performance								Statistical significance
		Accelerated		Within normal limits		Mildly delayed		Significantly delayed		
		N	%	N	%	N	%	N	%	
Motor 12th month	Boys	**11**	*5.6*	**168**	*85.7*	**15**	*7.6*	**2**	*1.0*	**ns**[a]
	Girls	**13**	*6.8*	**156**	*81.7*	**21**	*11.0*	**1**	*0.5*	
Motor 24th month	Boys	**7**	*3.6*	**167**	*87.0*	**18**	*9.*	**0**	-	p = 0.05
	Girls	**18**	*9.6*	**157**	*83.5*	**13**	*6.9*	**0**	-	
Motor 36th month	Boys	**23**	*12.5*	**155**	*84.2*	**6**	*3.3*	**0**	-	**p = 0.01**
	Girls	**40**	*21.7*	**142**	*77.2*	**1**	*0.5*	**1**	*0.5*	
Mental 12th month	Boys	**16**	*8.2*	**164**	*83.7*	**15**	*7.6*	**1**	*0.5*	**ns**
	Girls	**24**	*12.6*	**154**	*80.6*	**12**	*6.3*	**1**	*0.5*	
Mental 24th month	Boys	**22**	*11.4*	**149**	*77.2*	**21**	*10.9*	**1**	*0.5*	**p < 0.001**
	Girls	**56**	*29.6*	**116**	*61.4*	**17**	*9.0*	**0**	-	
Mental 36th month	Boys	**17**	*8.9*	**163**	*85.3*	**10**	*5.2*	**1**	*0.5*	**p = 0.007**
	Girls	**36**	*19.0*	**150**	*79.0*	**4**	*2.1*	**0**	-	

[a] ns - non-significant

Table II. Area under the ROC Curves with 95% confidence intervals for delayed performances of Bayley Scales' scores and early development milestones reported by mothers of 3-year-old children

		Bayley Motor Scale			Bayley Mental Scale		
		12th month	24th month	36th month	12th month	24th month	36th month
Lifting head while prone	ROC	*0.605*	0.591	0.637	*0.5255*	*0.5157*	0.397
	95%CI	0.554 - 0.654	0.539 - 0.641	0.585 - 0.686	0.474 - 0.577	*0.464 - 0.567*	0.347 - 0.449
Sitting without assistance	ROC	0.631	*0.549*	0.650	0.535	0.598	0.662
	95%CI	0.580 - 0.679	0.497 - 0.600	0.598 - 0.699	0.483 - 0.586	*0.546 - 0.648*	0.612 - 0.710
Standing without assistance	ROC	0.747	*0.606*	*0.517*	0.557	*0.511*	*0.510*
	95%CI	0.700 - 0.790	0.555 - 0.656	0.464 - 0.569	0.505 - 0.607	*0.459 - 0.563*	0.458 - 0.561
Walking alone	ROC	0.825	0.704	0.628	0.631	0.662	0.581
	95%CI	0.783 - 0.862	0.655 - 0.750	0.576 - 0.678	0.581 - 0.680	0.611 - 0.709	0.529 - 0.632
Walking upstairs	ROC	0.699	0.715	0.610	*0.536*	0.577	0.573
	95%CI	0.650 - 0.744	0.666 - 0.760	0.557 - 0.660	0.484 - 0.586	0.525 - 0.628	0.521 - 0.624
Bladder control	ROC	0.615	0.694	0.688	0.652	0.649	0.725
	95%CI	0.564 - 0.664	0.645 - 0.740	0.638 - 0.735	0.602 - 0.700	0.598 - 0..637	0.677 - 0.770
Bowel control	ROC	0.639	0.663	0.788	0.700	0.594	0.703
	95%CI	0.589 - 0.687	0.612 - 0.710	0.743 - 0.829	0.651 - 0.741	0.542 - 0.644	0.654 - 0.748
Speaking first meaningful words	ROC	0.605	0.629	0.748	0.603	0.610	0.606
	95%CI	0.553 - 0.654	0.578 - 0.679	0.699 - 0.792	0.551 - 0.652	0.559 - 0.660	0.555 - 0.654

ROC - Area under the ROC curve
95%CI - 95% Confidence Interval

significant results in 36th month while standing without assistance in 12th month. The biggest area under the ROC curve was obtained for walking alone and score on the motor scale of the BSID-II at 12th month - 0.825 and for bowel control and score on the motor scale of the BSID-II at 36th month - 0.788 (tab. II).

Sensitivity for the developmental milestones compared with score on the motor and mental scales of the BSID-II varied from 25.0% to 75.0%, specificity from 54.1% to 80.2%. Sensitivity most commonly varied about 60.0% with a specificity of about 70.0%. The best results were obtained for walking alone later than in the 12th month as a predictor of lower score on the motor scale of the BSID-II in 12th month. Because of the rather low value of specificity, the positive predictive ratios were about 10.0% with a maximum of 28.4%. The all of analysed milestones were characterized by low positive predictive value and rather high the negative one (tab. III).

There were no differences in accuracy of milestone assessment according to child gender, birth weight, duration of pregnancy, mother's age, educational level and whether mother take all-day care of child.

DISCUSION

To assess the concurrent validity of parent report developmental milestones, we tried to determine their diagnostic use and produced a cut-off time for classifying children's development. Because of moderate values of sensitivity and specificity, our parent report milestones demonstrated poor diagnostic utility in discriminating between children who have delayed development and those who were functioning within the normal range of the BSID-II scores. The high negative predictive values shown for developmental milestones as compared with the scores of the BSID-II tests indicate that clinical utility of this information could be useful to discriminate between children who attained developmental milestones fast as a group of low risk of developmental delays. This is not, however, a good screening tool to detect developmental delays (8). Pediatricians using the parent report developmental milestones should know the limitation of this method in the early detection of developmental problems (9). This is only a good prescreening procedure to detect children who attained the developmental milestones later than in the normal range, as a group which requires comprehensive, standardized screening tests in each of the major streams of development (10).

Only a few children in our group, drawn from the general population, showed very low development outcomes, then we decided to bring significantly and mildly delayed children together into "delayed" group. That group consists 2-22 children, dependent on year and studied scale (psychomotor or mental). The small number of milestones included in our analysis is a limitation of our study but it gave a chance to decrease a recall bias. Our previous analysis demonstrated that maternal reports of developmental milestones of children under 3 years old are reliable to be used in clinical judgment based on parental concern (7).

Table III. Sensitivity, specificity, negative and positive predicted values of developmental milestones based on cut off points

The age of milestone attainment		Bayley Motor Scale			Bayley Mental Scale		
		12th month	24th month	36th month	12th month	24th month	36th month
Lifting head while prone later than in the 2nd month	Sens.	62.5 24.7-91.0	50.0 31.3-68.7	54.1 36.9-70.5	42.9 24.5-62.8	42.1 26.3-59.2	26.7 7.8-55.1
	Spec.	61.3 56,0-66.4	60.5 55.1-65.7	61.3 56.0-66.5	60.1 54.7-65.2	59.8 54.3-65.0	59.4 54.2-64.6
	+ PV	3.5 1,2-8,0	9.9 5.7-15.8	13.1 8.2-19.5	7.8 4.1-13.3	10.5 6.1-16.5	2.7 0.7-6.7
	- PV	98.6 96.1-99.7	93.3 89.1-96.2	92.5 88.3-95.6	93.0 88.9-95.9	90.2 85.5-93.7	95.1 91.4-97.2
Sitting without assistance later than in the 6th month	Sens.	64.9 47.5-79.8	46.7 28.4-65.7	25.0 3.9-65.0	50.0 30.7-69.3	57.9 40.8-73.7	73.3 44.9-92.0
	Spec.	57.7 50.6-61.2	56.0 50.5-61.3	66.9 61.7-71.7	56.0 50.6-61.2	57.0 515.68-3.	56.5 51.2-61.7
	+ PV	14.2 9.3-20.3	8.5 4.7-13.8	1.7 0.2-5.9	8.3 4.6-13.5	13.2 8.4-19.3	6.6 3.33-11.5
	- PV	93.8 89.7-96.7	92.3 87.8-95.5	97.5 94.7-99.1	93.4 84.2-92.9	92. 87.8-95.5	98.1 95.1-99.5
Standing with assistance later than in the 9th month	Sens.	67.6 50.2-82.0	40.0 22.7-59.4	25.0 3.9-65.0	42.9 24.5-62.8	28.9 15.4-45.9	26.7 8.0-55.1
	Spec.	70.0 64.8-74.8	67.9 62.7-72.8	66.9 61.7-71.7	67.0 61.9-17.9	66.5 61.1-71.5	66.0 60.9-70.9
	+ PV	19.5 13.1-27.5	9.8 5.1-16.5	1.7 0.2-5.9	9.4 4.9-15.8	8.9 4.5-15.3	3.2 0.9-7.9
	- PV	95.2 91.8-97.5	92.8 88.9-95.7	97.5 94.7-99.1	93.7 89.9-96.3	89.2 84.7-92.8	95.6 92.2-97.8
Walking alone later than in the 12th month	Sens.	73.0 55.9-86.2	56.7 37.4-74.5	37.5 9.0-75.3	46.4 27.5-66.1	42.1 26.3-59.2	33.3 11.9-61.6
	Spec.	80.2 75.6-84.3	78.5 73.8-82.7	76.8 72.1-81.1	76.8 72.0-81.1	77.5 72.7-81.9	75.8 71.1-80.2
	+ PV	28.4 19.6-38.6	18.7 11.3-28.2	3.5 0.7-10.0	13.7 7.5-22.3	17.4 10,3-26.7	5.4 1.8-12.2
	- PV	96.5 93.7-98.3	95.4 92.3-97.5	98.2 95.8-99.4	94.8 91.5-97.0	92.3 88.5-95.1	96.5 93.6-98.3
Walking upstairs later than in the 18th month	Sens.	56.8 39.5-72.9	60.0 40.6-77.3	37.5 9.0-75.3	35.7 18.7-55.9	39.5 24.1-56.6	33.3 11.9-61.6
	Spec.	75.0 70.1-79.5	75.3 70.4-79.8	72.9 67.9-77.4	72.5 67.5-77.1	73.7 68.6-78.3	72.2 67.3-76.8
	+ PV	19.6 12.6-28.4	17.5 10.7-26.2	3.0 0.6-8.6	9.3 4.6-16.5	14.4 8.3-22.7	4.8 1.6-10.8
	- PV	94.2 90.7-96.6	95.6 92.4-97.7	98.1 95.6-99.4	93.4 89.8-96.1	91.5 87.6-94.6	96.3 93.3-98.2
Bladder control later than in the 33rd month	Sens.	41.0 25.6-51.9	51.6 33.1-69.8	75.0 35.0-96.1	48.3 29.5-67.5	51.3 24.8-67.6	66.7 38.4-88.1
	Spec.	73.3 68.2-61.7	73.6 68.6-78.2	73.0 68.1-77.6	73.4 68.5-78.0	74.3 69.3-78.9	73.5 68.6-78.0
	+ PV	14.8 8.7-22.9	15.0 8.9-23.1	5.9 2,2-12,4	13.0 7.3-20.8	18.7 11,8-27,4	9.4 4,6-16,7
	- PV	91.6 87.7-94.6	94.4 91.1-96.8	99.2 97,3-99,9	94.5 91.2 -96.9	93.0 89,3-95,7	98.2 95,7-99,4
Bowel control later than in the 29th month	Sens.	56.4 39.6-72.2	51.6 33.1-69.8	75.0 35.0-96.1	62.1 42.3-79.3	56.4 39.6-72.2	66 7 38.4-88.1
	Spec.	73.9 68.9-78.5	72.8 67.8-77.4	72.5 67.6-77.1	73.5 68.6-78.0	74.1 69.1-78.7	72.5 67.5-77.0
	+ PV	19.6 12.7-28.2	14.5 8,5-22,5	5.8 2.1-12.1	16.1 9.8-24.2	20.0 13,0-28,7	9.1 4.4-16.1
	- PV	93.7 90.2-96.3	94.4 90.9-96.8	99.2 97.2-99.9	96.0 92.9-98.0	93.7 90,1-96,3	98.1 95.7-99.4
Speaking first meaningful words later than in the 18th month	Sens.	64.9 47.5-79.8	73.3 54.1-87.7	75.0 35.0-96.1	67.9 47.6-84.1	68.4 51.3-82.5	66.7 38.4-88.1
	Spec.	54.5 49.1-59.9	55.7 50.3-61.1	54.4 49.0-59.7	54.3 48.9-59.6	55.8 50.3-61.2	54.1 48.7-59.3
	+ PV	13.4 8.8-19.3	12.7 8.1-18.6	3.6 1.3-7.8	10.6 6.5-16.1	14.9 10,0-21,4	5.7 2,8-10,3
	- PV	93.5 89.1-96.5	96.0 92.2-98.2	99.0 96.3-99.9	95.5 91.6-97.9	94.0 89,7-96,8	97.5 94,2-99,2

Sens. – Sensitivity + PV - Positive predictive value
Spec. – Specificity - PV - Negative predicted value

CONCLUSION

Parent report developmental milestones are a better tool for excluding those children who attain milestones rapidly, as a group with low risk of developmental delays, than in identifying children whose development is suspected of being delayed. Our study has confirmed the recommendation for development surveillance in pediatric care that information from parents about infants should be combined with clinical observation and standardized developmental assessment.

ACKNOWLEDGEMENTS

This is a part of ongoing comparative longitudinal study on the health impact of prenatal exposure of infants and children to outdoor/indoor air pollution, which is being conducted in the New York City and Krakow. The study received funding from an RO1 grant entitled "Vulnerability of the Fetus/Infants to PAH, $PM_{2.5}$ and ETS" (RO1 ES10165 NIEHS; 02/01/00 - 01/31/04) and from NIEHS (5 RO1 ES010165-0451), the Lundin Foundation, and the Gladys T. and Roland Harriman Foundation. Principal Investigator: Prof. F.P. Perera, Co-investigator: Prof. W. Jedrychowski

REFERENCES

1 Oberklaid F, Efron D. Developmental delay--identification and management. Aust Fam Phys 2005;34(9):739-42.
2. Glascoe FP, Robertshaw NS. New AAP policy on detecting and addressing developmental and behavioral problems. J Pediatr Health Care 2007; 21(6):407-12.
3. Voigt RG, Brown FR III, Fraley JK, Llorente AM, Rozelie J, Turcich M et al. Concurrent and predictive validity of the Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS) and the Mental Developmental Index of the Bayley Scales of Infant Development. Clin Pediatr 2003; 42(5):427-436.
4. Harris S, Megens A, Backman C, Hayes V. Stability of the Bayley II Scale of Infant Development in a sample of low-risk and high-risk infants. Develop Med Child Neurol 2005;47:820-823.
5. Bayley N. Bayley Scales of Infant Development. New York: The Psychological Corporation,1969.
6. Jedrychowski W, Whyatt RM, Camann DE, Bawle UV, Peki K, Spengler JD, Dumyahn TS, Penar A, Perera FF. Effect of prenatal PAH exposure on birth outcomes and neurocognitive development in a cohort of newborns in Poland. Study design and preliminary ambient data. Int J Occup Med Environ Health 2003;16(1):21-9.
7. Majewska R, Mrożek-Budzyn D, Kiełtyka A, Augustyniak M. Usefulness of maternal assessment of children development based on reported age of acheived milestones. Przegl Epidemiol 2013; 67:487-490.
8. Taanila A, Murray GK, Jokelainen J, Isohanni M, Rantakallio P. Infant developmental milestones: a 31-year follow-up. Dev Med Child Neurol 2005;47(9):581-6.
9. Noyes-Grosser DM, Holland JP, Lyons D, Holland CL, Romanczyk RG, Gillis JM. Rationale and methodology for developing guidelines for early intervention services for young children with developmental disabilities. Inf Young Children;2005: 18, 119-135.
10. Glascoe FP. Parents' concerns about children's development: prescreening technique or screening test? Pediatrics 1997;99(4):522-8.

Received: 25.03.2013
Accepted for publication: 21.01.2014

Address for correspondence:
Dorota Mrożek-Budzyn
Department of Epidemiology and Preventive Medicine
Jagiellonian University - Medical College
ul. Kopernika 7a, 31-034 Kraków
tel./fax: +48 12 4231003 / +48 12 4228795
e-mail. dorota.mrozek-budzyn@uj.edu.pl