

Validity and Reliability Determination of Denver Developmental Screening Test-II in 0-6 Year-Olds in Tehran

Soheila Shahshahani¹, MD; Roshanak Vameghi², MD; Nadia Azari¹, MD; Firoozeh Sajedi^{*2}, MD; and Anooshirvan Kazemnejad³, PhD

1. Pediatric Neurorehabilitation Research Center, University of Social Welfare & Rehabilitation Sciences, Tehran, IR Iran
2. Clinical sciences department, Pediatric neurorehabilitation research center, USWR, Tehran, IR Iran
3. University of Tarbiat Modares, Tehran, IR Iran

Received: Jun 13, 2009; Final Revision: Jan 07, 2010; Accepted: Feb 28, 2010

Abstract

Objective: This research was designed to identify the validity and reliability of Persian version of Denver Developmental Screening Test II (DDST-II) in Iranian children, in order to provide an appropriate developmental screening tool for Iranian child health workers.

Methods: At first a precise translation of test was done by three specialists in English literature and then it was revised by three pediatricians familiar with developmental domains. Then, DDST-II was performed on 221 children ranging from 0 to 6 years, in four Child Health Clinics, in north, south, east and west regions of Tehran to determine the agreement coefficient, these children were also evaluated by ASQ test. Because ASQ is designed to use for 4-60 month- old children, children who were out of this range were evaluated by developmental pediatricians. Available sampling was used. Obtained data was analyzed by SPSS software.

Findings: Developmental disorders were observed in 34% of children who were examined by DDST-II, and in 12% of children who were examined by ASQ test. The estimated consistency coefficient between DDST-II and ASQ was 0.21, which is weak. DDST-II and the physicians' examination was 0.44. The content validity of DDST-II was verified by reviewing books and journals, and by specialists' opinions. All of the questions in DDST-II had appropriate content validity, and there was no need to change them. Test-retest and Inter-rater methods were used in order to determine reliability of the test, by Kuder-Richardson's α and Kuder-Richardson coefficients. Kuder-Richardson coefficient for different developmental domains was between 61% and 74%, which is good. Cronbach's α coefficient and Kuder-Richardson coefficient for test-retest were 92% and 87% and for Inter-rater 90% and 76%, respectively.

Conclusion: This research showed that Persian version of DDST-II has a good validity and reliability, and can be used as a screening tool for developmental screening of children in Tehran city.

Iranian Journal of Pediatrics, Volume 20 (Number 3), September 2010, Pages: 313-322

Key Words: Development; Developmental screening; DDST-II

* Corresponding Author;

Address: Clinical sciences department, University of Social Welfare & Rehabilitation Sciences, Tehran, IR Iran

E-mail: fisajedi@uswr.ac.ir

Introduction

Developmental disabilities can be seen in 10-15% of children in different populations. Early detection and appropriate referral of children with developmental delays or disorders is important in Pediatrics. This is only possible by continuous developmental monitoring and assessment. Developmental assessment is made by early detection of problems through developmental surveillance and screening, precise evaluation by using standardized and formal diagnostic tools as well as evaluation of the medical, social, family history and physical examination of the child^[1,2]. Developmental screening must be repeated periodically and incorporated into pediatrics practice^[3,4].

Developmental screening test is a brief standardized tool that is used for identifying children who need more detailed evaluation^[5] and if used appropriately is useful and cost benefit effective^[6]. Because screening is used for identifying the children who will receive the benefits of more professional evaluation or treatment, it is recommended that all children be screened for developmental delays^[5].

There are many developmental screening tools. The base of all of them is achieving developmental milestones at specific chronological ages. Denver Developmental Screening Test II (DDST-II) and Bayley are examples for such formal tools. For having ability to differentiate between abnormal children from those normal children who have slower rate of achieving developmental skills, these developmental screening tools must be reliable and valid, have acceptable sensitivity and specificity, be easy to perform and not expensive^[1,6,7].

DDST-II is a formal developmental screening tool that assesses children from birth to 6 years of age. First it was standardized on 1036 children (543 boys and 493 girls) from 2 weeks old to 6/4 years of age in Denver, Colorado as DDST^[8]. Then in 1992 it is revised and restandardized on 2096 children and is known as DDST-II. Test reliability on test-retest is 90% and its inter-rater reliability is 80-95%^[9]. The test is valid and there is a strong relationship between classification on the DDST and scores on the Stanford-Binet intelligence scales and the

Previous edition of Bayley infant scales^[10].

DDST-II is a brief and validated screening tool that many of pediatricians are familiar with it. Although there is doubt about its limited specificity (43%) and risks of over referral^[5,11], it has high rate of sensitivity (83%) and identifies children with developmental delays^[5,12]. DDST-II assesses child's development in 4 general areas: 1) personal-social (25 items), 2) fine motor- adaptive (29 items), 3) language (39 items), and 4) gross motor (32 items)^[9,13]. Screening by it produces 3 scores: normal, suspect and untestable^[9] (these children refused participating in some items that 95% of age matched children could pass them). Sometimes DDST results are interpreted as normal, suspect, questionable (these children cannot pass some items that 75-95% of age matched children could pass them) and untestable. A study found sensitivity of 80% if "questionable scores" were included with abnormal scores but specificity of 46%. Alternatively, if "questionable scores" were included with normal scores, sensitivity was 46% and specificity 80%^[11].

By considering the importance of early detection of developmental disabilities and absence of an Iranian developmental screening test, this study was planned to determine the validity and reliability of Persian version of DDST-II (by translating to Persian and evaluating the cultural adaptation of the items) in Iranian children in order to provide an appropriate developmental screening tool for Iranian child health workers.

Subjects and Methods

This research is an action research that was performed from January to August 2008 in 4 Child Health Care centers located in north, south, east and west regions of Tehran city.

These are primary health care centers which provide mainly general health services for people including children from different socio-economical classes of general population. Usually normal children visit such centers and services for growth monitoring, vaccination, vitamin supplements, etc). These centers are

under the supervision of Deputy of Health, or Shahid Beheshti Medical University.

At first test form and guiding sheet was translated precisely by 3 specialists familiar with English. Then the research team (4 pediatricians) read all 3 translated versions and for each item in form and sheet we chose the best translation (simple, short, easy to understand and culturally compatible). Then we sent them along with original version to 3 other pediatricians who were familiar with developmental domains. The research team discussed their view points and implemented their opinions in the final form. Healthy newborns, infants and children, 0-6 years old, in Tehran city could participate in this study. The inclusion criteria were: 1) age between birth to 6 years, 2) Iranian nationality, 3) living in Tehran city, and 4) parental cooperation. Exclusion criteria were: 1) having obvious developmental delay or disability (because including children with gross developmental disorders would lower the cutoff point for each developmental item in Iranian children), 2) parental refusal.

The study was approved by the research committee and thereafter by the ethical committee of University of Social Welfare & Rehabilitation Sciences. Consent for participation was obtained from parents. The parents whose children had developmental problems were informed and guided.

Convenient sampling was used and 221 children (100 girls and 121 boys) in 13 age groups (0 to 2, 2.1 to 4, 4.1 to 6, 6.1 to 9, 9.1 to 12, 12.1 to 15, 15.1 to 18, 18.1 to 24, 24.1 to 30, 30.1 to 36, 36.1 to 48, 48.1 to 60 and 60.1 to 72 months), each age group containing 17 children, were examined. Demographic items included date of birth, sex, birth order, maternal education level, gestational age at birth (preterm or term; for preterm children up to 2 years we calculated corrected age), and history of disability of the child.

Eight examiners were trained in a 1 day workshop for performing the DDST-II. A demographic questionnaire was completed for each child by parents and then DDST-II was done by the examiners. In order to determine the reliability of DDST-II, 25% of children in each age group (small children after 30-60 minutes,

older children up to 2-3 days later) were re-examined by the same examiners (test-retest).

Another 25% of children were retested by another examiner (inter-rater reliability). In order to determine agreement coefficient, these children were also evaluated by ASQ (Ages and Stages Questionnaires) test. ASQ is not a diagnostic gold standard test. It is a developmental screening tool. Because we had no accessibility to any diagnostic tests we compared these two developmental screening tools to determine their agreement coefficient. Anyway, by another research team, ASQ was translated into Persian and was standardized on 11000 Iranian children. The results have not been published yet, but the general report exists and we have used the translated forms. Because ASQ is designed to use for 4-60 month-old children and each questionnaire can be used for one month before or after the specific age, children who were out of this range (3-61 months) were evaluated by developmental pediatricians. Also 10% of children of other age groups, after examining by DDST-II and completing the ASQ by parents, were evaluated by developmental pediatricians.

As mentioned above, test-retest and inter-rater methods were used in order to determine reliability of the test by Cronbach's α and Kuder-Richardson coefficients. We use Cronbach's α for reliability determining of each test item and Kuder-Richardson coefficient in 4 developmental domains. In test-retest and inter-rater tests we measured Cronbach's α and kappa measure of agreement for comparison of each developmental domain and final results of each test respectively. Content validity of the test was verified by reviewing texts and related articles, and by specialists' opinions. Data was analyzed by SPSS software.

Findings

In this study 221 children were evaluated by DDST-II (100 girls and 121 boys) in 13 age groups (Appendix). Birth order of children were 73% first, 22% second, 3% third and 2% fourth

Table 1: Interclass Correlation Coefficients for test – retest examination by DDST- II

Developmental domain	ICC*	ICC interval (95% confidence interval)	Cronbach's α coefficient
Personal-social	0.95	0.92-0.97	0.96
Fine motor-adaptive	0.90	0.84	0.95
Language	0.93	0.89-0.96	0.96
Gross motor	0.91	0.86-0.95	0.95

ICC: Interclass Correlation Coefficients/ DDST-II: Denver Developmental Screening Test II

child of family. Maternal educations of 85% of children were at high school or greater level. 95% of children were born at term and 5% of them preterm (for preterm children up to 2 years we calculated and considered corrected age).

Children were selected from 4 different regions of Tehran city. Developmental screening of children by DDST-II showed that 143(65%) of them developed normally, 75(34%) had developmental delay (suspect) and 3(1%) were untestable according to test scoring method.

Cautions and delays number in each developmental domains are 13 and 20 in Personal-social, 13 and 24 in Fine motor-adaptive, 21 and 16 in language and finally 10 and 23 in Gross motor areas. As it is seen number of cautions and delays are greater in language and fine motor- adaptive domains respectively. Children with developmental delays differed in number of affected domains. 36 children had delay in 1, 27 children in 2 and 9 children in 3 developmental domains.

In this study, reliability was evaluated by the Kuder-Richardson coefficients determination. The estimated coefficients were 0.74 for personal-social, 0.63 for fine motor-adaptive, 0.63 for language and 0.61 for gross motor domains. Test-retest and inter-rater methods were also used as other ways for reliability determination. Interclass Correlation

Coefficients for test-retest and inter-rater methods are shown in Tables 1 and 2.

Table 3 shows comparison between the results of DDST-II with ASQ and results of pediatricians' evaluation. Comparison of DDST-II and ASQ results showed that 109 children passed two tests and 21 children failed in both of them.

Consistency coefficient between DDST-II and ASQ was 0.21. Thus sensitivity and specificity of DDST-II could be calculated as shown below:

$$\text{Sensitivity} = 21: 35 \times 100 = 60\%$$

$$\text{Specificity} = 109:158 \times 100 = 69\%$$

Comparison of DDST-II and results of pediatricians' evaluation showed that 42 children passed and 4 children failed in both evaluations. Consistency coefficient between DDST-II and pediatricians' evaluation was 0.44. Final translated version of DDST-II has been shown in appendix.

Discussion

In our study the content validity of DDST-II was verified by reviewing books and journals, and by specialists' opinions. All of the questions in DDST-II had appropriate content validity, and there was no need to change them.

Table 2: Interclass Correlation Coefficients for inter-rater examination by DDST-II

Developmental domain	ICC*	ICC interval (95% confidence interval)	Cronbach's α coefficient
Personal-social	0.88	0.79-0.93	0.94
Fine motor-adaptive	0.89	0.82-0.94	0.94
Language	0.96	0.93-0.98	0.98
Gross motor	0.86	0.77-0.92	0.93

ICC: Interclass Correlation Coefficients/ DDST-II: Denver Developmental Screening Test II

Table 3: Comparison between the results of DDST-II with ASQ and pediatricians' evaluation

Test or clinical evaluation	Result	DDST-II results (No. of children)		Total
		Failed	Passed	
ASQ (No. of children)	Failed	21	14	35
	Passed	49	109	158
	Total	70	123	193
Pediatricians' evaluation (No. of children)	Failed	4	0	4
	Passed	10	42	52
	Total	14	42	56

DDST-II: Denver Developmental Screening Test II / ASQ: Ages and Stages Questionnaires

We evaluated the reliability of the test by the Kauder-Richardson coefficients determination. Determined Kauder-Richardson coefficients for all of developmental domains were "good". Test-retest and inter-rater methods were also used as other ways for reliability determination. In test-retest examination the Cronbach's α coefficients for all developmental domain is very good and kappa measure of agreement is 87% ($P < 0.001$). In Inter-rater examination the Cronbach's α coefficients for all developmental domain were very good and kappa measure of agreement was 76% ($P < 0.001$). Thus DDST-II has very good reliability in test-retest and Inter-rater examination.

Sensitivity of DDST-II in different references ranges from 40-83%^[14,15] and its specificity is reported from 40-80%^[16,17]. In this study sensitivity and specificity of DDST-II by comparing the results of DDST-II and ASQ were 60% and 69% respectively. Of course this cannot be considered as the actual validity of the test, because, as explained before, ASQ is not a diagnostic gold standard test. We found that children passed the ASQ (88%) more than DDST-II (65%) and consistency coefficient of the two tests was poor (0.21). Therefore, either ASQ may be undersensitive and/or DDST-II oversensitive. Which of these is true? It has to be investigated by comparing the results of these tests with the results of a developmental diagnostic test.

It is possible that in comparison with the Denver sample, Iranian children have a slower rate of development. One study conducted in Shiraz (Iran) showed that 3-6 year-old Iranian children have slower rate of development by DDST-II in fine and gross motor domains^[16]. It is worthy to mention that DDST-II has subgroup standards based on sex, race, maternal education

and place of residence that are presented in Denver-II technical manual^[18].

Some of our other findings revealed no relationships between sex, maternal education and place of residence with children's developmental status in Tehran. In another native study in Shiraz city, gross and fine motor performance of 1524 children aged 3-6 years was evaluated by DDST-II in 2005-2008. In this evaluation girls had better performance^[16]. Yalaz and Epir tested 1176 Turkish children aged 2 weeks to 6 yr-4 mo by DDST in 1984. They found that girls' development is better than that of boys'^[17].

Durmazlar and Anlar also evaluated 1091 Turkish children aged 0-72 months by DDST-II in 1998. In their study few and inconsistent differences were observed between boys and girls^[19]. Bryants and Stark. evaluated the achievement of test items in the first year of life of Cardiff infants by DDST in 1974. They concluded that in the first year of life, there are no developmental differences between boys and girls^[20].

In this study maternal education had no effect on children's developmental status. In South Okanogan of Canada, Barnes and Stark evaluated 206 children 2 weeks to 6 yr-4 mo old by DDST. In their study, there was no relationship between maternal education and child's developmental level except in 10-12 month-old infants^[21]. William's study in Philippine showed that maternal education and her birth place (rural or urban) had relation to DDST results. Children of mothers with higher educational level and who were born in urban regions had better results^[22]. Lejjaraga and co-workers studied 0-5 year-old children in Argentina in 2002. In their research, after first year of life, sex

and maternal education were related to children's developmental status^[23].

Bryant et al examined 686 infants aged 2 weeks to 12 months in Cardiff by DDST in 1974. They found that children of Cardiff in fine and gross motor domains had slower developmental rate^[20]. Later they standardized DDST on 1574 under 6 year-old children in Cardiff in 1976.

They showed that Cardiff children had slower developmental rate, but under 18 months of age, they were better than Denver children in language domain. There was no difference between the two groups in fine motor development and Cardiff children had better performance in personal-social domain^[24]. Oeda in one DDST study on 1171 children showed that children of Tokyo had better results in some items of personal-social domain but in infancy they had slower developmental rate in gross and fine motor domains^[25]. Gross and fine motor performance of 78 healthy Swedish children aged 15 to 18 months were examined with DDST by Lundberg. He showed that Swedish children had slower rate of fine and gross development^[26]. Shapiro and Harles in Israel examined 2248 children 2 weeks to 6.5 years old with DDST. They found that Israeli children in comparison to Denver children had slower rate of development in gross and fine motor domains^[27]. A review article showed that by using DDST, children of Japan, Philippine, Tokyo, Okinawa, Netherland and Bangkok had slower rate of motor development^[28].

This study has some limitations. First, some parents that their children had to be re-examined by DDST-II, did not return to clinic. So we chose other cases and there was wasting of time and resources. Second, developmental screening tools are not diagnostic and their results must be followed by a more intensive evaluation. The sensitivity and specificity of DDST-II must be determined by comparison of the test results with a developmental diagnostic test. Because there was no standardized diagnostic test in Iran, we compared the DDST-II with ASQ and pediatricians' evaluation of children's development. Third, by considering the results of similar researches, and on the base of results of this study, delays in fine and gross motor areas are more than in other

developmental domains. It is recommended to design another study on larger samples in order to standardized DDST-II and determine the norms of Iranian children.

Early detection and intervention of children with developmental delays or disorders is an important issue in pediatrics medicine. Early detection and intervention in developmental problems can reduce their impacts on the well-being and functioning of child and his/her family. American Academy of Pediatrics recommended that pediatricians use standardized developmental screening test regularly at the 9, 18 and 30 (or 24) month visits^[29].

Conclusion

This research showed that Persian version of DDST-II has a good validity and reliability, and can be used as a screening tool for developmental screening of children in Tehran city. For determining the sensitivity and specificity of the test, it is suggested that the results of each of the two screening tests DDST-II and ASQ are compared with a standard diagnostic test in future studies.

Acknowledgment

This research was supported by a grant from Pediatric Neurorehabilitation Research Center of University of Social Welfare and Rehabilitation, Tehran. We wish to thank Dr N Hatamizadeh, Dr H Karimi and Dr F Soleimani for their assistance.

Conflict of Interest: None

References

1. Glasco FP. Developmental screening & surveillance. In: Kliegman RM, Behrman RE, Jenson HB, Stanton BM (eds). Nelson Textbook of Pediatrics. 18th ed. Philadelphia: Saunders. 2008; Pp: 74-81.

2. Glascoe FP. Screening for developmental and behavioral problems. *Ment Retard Dev Disabil Res Rev.* 2005;11(3):173-9.
3. Hix-Small H, Marks K, Squires J, et al. Impact of implementing developmental screening at 12 and 24 months in a pediatric practice. *Pediatrics.* 2007;120(2):381-9.
4. Earls MF, Hay SS. Setting the stage for success: Implementation of developmental and behavioral screening and surveillance in primary care practice - The North Carolina Assuring Better Child Health and Development (ABCD) Project. *Pediatrics.* 2006;118(1):e183-8.
5. Screening infants and young children for developmental disabilities. American Academy of Pediatrics Committee on Children with Disabilities. *Pediatrics.* 1994;93(5):863-5.
6. Levine DA. Guiding parents through behavioral issues affecting their child's health: the primary care provider's role. *Ethn Dis.* 2006;16(2 Suppl 3):S3-21-8.
7. Frankenburg WK. Developmental surveillance and screening of infants and young children. *Pediatrics.* 2002;109(1):144-5.
8. Frankenburg WK, Dodds JB. The Denver developmental screening test. *J Pediatr.* 1967; 71(2):181-91.
9. Frankenburg WF, Dodds J, et al. DENVER II Training Manual. 5th ed. Denver, Colorado/ Denver developmental material Inc.1992. Pp 4.
10. Salvia J, Ysseldyke JE. Assessment in special and inclusive education. 9th ed. Boston; Houghton Mifflin. 2004; Pp: 46-9.
11. Ahsan S, Murphy G, Kealy S, et al. Current developmental surveillance: Is it time for change? *Ir Med J.* 2008; 101(4):110-2.
12. Glasco FP. Are overreferrals on developmental screening tests really a problem? *Arch Pediatr Adolesc Med.* 2001;155(1):54-9.
13. Robert D. Needlman. Developmental assessment. In: Kliegman RM, Behrman RE, Jenson HB, Stanton BM (editors). *Nelson Textbook of Pediatrics.* 17th ed. Philadelphia: Saunders. 2004; p62-6
14. American Academy of Pediatrics, Council on Children with Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children with Special Needs Project Advisory Committee, Identifying Infants and Young Children With Developmental Disorders in the Medical Home: An algorithm for developmental surveillance and screening. *Pediatrics.* 2006;118(1):405-20.
15. Vameghi R, Sajedi F, Shahshahani S, Hatamizadeh N. Early detection, diagnosis and introduction to early intervention in childhood developmental problems. First ed. Tehran: University of Social Welfare and Rehabilitation Sciences. 2005; 35-53. (in Persian)
16. Pasand F. Standardization and validity and reliability determination of DDST-II for fine and gross motor function of 3-6 year old children in Shiraz city, dissertation for MS. Tehran, Tarbiat Modares University, 2008.(in Persian)
17. Yalaz K, Epir S. Urban Turkish children performance on the Denver Developmental Screening Test. *Develop Med Child Neurol.* 1984;26(5):632-43.
18. Frankenburg WK, Dodds J, et al, DENVER II Technical Manual. 2th ed. Denver, Colorado/ Denver developmental material Inc. (1996)Pp 55-59.
19. Durmazlar N, Anlar B. Turkish children performance on Denver Developmental Screening Test II: Effect of sex and mothers education; *Develop Med Child Neurol.* 1998; 40: 411-6.
20. Bryant GM, Davies KJ, Newcombe RG. The Denver Developmental Screening Test, Achievement of test items in the first year of life by Denver and Cardiff infants. *J Develop Med Child Neurol.* 1974; 16(4):475-84.
21. Barnes KE, Stark A. The Denver Developmental Screening Test, A Normative Study. *Am J Public Health.* 1975; 65(4):363-9.
22. William PD. The Metro-Manila. The Developmental Screening Test; A Normative Study, *Nursing Research.* 1984; 33(4):208-12.
23. Lejarraga H, Pascucci MC, Krupitzky S, et al. Psychomotor development in Argentinean children aged 0-5 years. *Paediatr Perinat Epidemiol.* 2002;16(1):47-60.
24. Bryant GM, Davies KJ, Newcombe RG. Standardization of the Denver Developmental Screening Test for Cardiff children. *Dev Med Child Neurol.* 1979;21(3):353-64.
25. Ueda RA. Standardization of the Denver Developmental Screening Test on Tokyo children. *Dev Med Child Neurol.* 1978;20(5): 647-56.
26. Lundberg A. Gross and fine motor performance in healthy Swedish children aged fifteen and eighteen months. *Neuropadiatrie.* 1979;10(1): 35-50.
27. Shapira Y, Harel S. Standardization of the Denver Developmental Screening Test for Israeli children. *Isr J Med Sci.* 1983;19(3):246-51.
28. Mayson TA, Harris SA, Bachman CL. *Pediatric Physical Therapy;* Summer 2007;19(2), pp 148-52.
29. American Academy of Pediatrics: Developmental surveillance and screening of infants and young children. *Pediatrics.* 2001; 108(1):192-5.