A Study on Early Prediction System for Developmental Delays

著者	MURAI NORIO, ADACHI TOMOAKI, SHIGEMASU KAZUO, NIHEI YOSHIAKI, MURAI NORIKO, TANNO KANEO
journal or publication title	Tohoku psychologica folia
volume	50
page range	15-24
year	1992-05-01
URL	http://hdl.handle.net/10097/62525

Tohoku Psychologica Folia
1991, **50**, 15-24

A STUDY ON EARLY PREDICTION SYSTEM FOR DEVELOPMENTAL DELAYS

By

MURAI Norio (村井憲男)[1], ADACHI Tomoaki (足立智昭),
(*Tohoku University*) (*Miyagi Gakuin Women's College*)
SHIGEMASU Kazuo (繁桝算男), NIHEI Yoshiaki (仁平義明),
(*Tokyo Institute of Technology*) (*Tohoku University*)
MURAI Noriko (村井則子) and TANNO Kaneo (丹野兼夫)
(*Tohoku Social Welfare University*) (*Tanno Pediatric Clinic*)

The present study aimed at confirming the effectiveness of our early prediction system for developmental delays which was a sort of expert system applying the Dempster-Shafer's theory. The subjects were 210 infants of 18 months of age who were drawn from the sample population of 2077 infants born at Natori City, 53 of them being developmentally delayed and 157 of them being not delayed. On the basis of the early predictors during the first six months which included information obtained from medical records and a maternal questionnaire, our prediction system could discriminate between delayed and non-delayed infants of 18 months with overall agreement at 80 percent or more. In further analyses of cross-validation the high discriminability of the present system was confirmed. These findings suggest that our prediction system is applicable to general practical use.

Key words: developmental delays, early screening, infants, expert systems, Dempster-Shafer's theory, ignorances.

Research on developmental medicine and psychology for nearly three decades has traced the relationship between various events in early infancy and later developmental delays. These events as early predictors for later developmental delays include such events as perinatal anoxia, prematurity, low birth weight, socioeconomic status and so on. Although the correlations between individual predictors and later delays have become more and more clear and detailed by recent studies, it is a common statement that the prediction systems based on those predictors have ineffective predictability for subsequent developmental problems (e.g., Broman, 1979; Bronstein & Sigman, 1986; Sigman & Parmelee, 1979).

The cause of the ineffective predictability might be attributed in part to analytic techniques of data. In most studies on early prediction of development, their predic-

1. Correspondence concerning this article should be addressed to Norio Murai, Department of Developmental Disorders, Faculty of Education, Tohoku University, Kawauchi, Aoba-ku, Sendai 980, Japan.

tion formulae for identifying developmental delays were derived from statistical and mechanical analyses such as cumulative scores of risk factors (e.g., Parmelee, Kopp, & Sigman, 1976) and multiple discriminant analysis (e.g., Field et al., 1978). The discrimination rule statistically selected as optimal to a particular sample is not always to be true for another sample population.

Our previous study (Murai et al., 1990) attempted to predict developmental delays at 18 months depending on the information obtained from mothers' observation of their 6-month infants. In this study a diagnostic system applying the Dempster-Shafer's theory (Shafer, 1976) was used as a prediction system. The system, which can be considered as a sort of expert system utilizing professional knowledge, effectively discriminated the delayed infants at 18 months from the non-delayed : nearly 80% of the subjects were correctly identified. The cross-validity of the discrimination rule adopted by the system was also confirmed. These findings suggest that the analytic technique of data is one of significant factors in constructing effective early prediction systems.

The main aim of the present study was to confirm further the effectiveness of our prediction system mentioned above : as compared with two prediction systems based on statistical and mechanical analyses, how effectively can our system predict the developmental delays at 18 months of age by the information obtained during the first six months of life ?

METHOD

Population studied : Subjects were 210 infants of 18 months of age in two groups, the delayed and the non-delayed groups. Out of 2077 infants who were born between September 1983 and July 1987 at Natori City in Miyagi Prefecture and visited the mother-infant health center at this district for their developmental checkup at 18 months of age, all those identified as developmentally delayed (having no known organic problem but having delays in motor and/or cognitive development) consisted of the delayed group, numbering 53 infants (39 males and 14 females). As the non-delayed group, 157 infants (121 males and 36 females, approximately three times the number of the delayed group, chosen on the basis of corresponding sexes and birth dates) were chosen from those having no definite developmental problem at 18 months.

Variables analyzed : Two types of variables were included in the present analysis. The first, variables based on developmental records, consisted of 28 variables including perinatal, neonatal and postnatal medical events (Table 1). Information was obtained from the developmental checkup records of the infants collected at three checkup visits, the newborn, 3- and 6-month-olds. The second, variables based on maternal information, consisted of 35 variables based on our questionnaire named as EASY (Checklist for Early Atypical Signs of the Young, Murai et al., 1986). The questionnaire consists of abnormal features of infants which their mothers would be

Table 1. List of variables based on developmental records

1)	Estimated gestational age (EGA)
2)	Type of delivery (TYPEDEL)
3)	Abnormalities of labor (ABNORML)
4)	Weight at birth (WB)
5)	Length at birth (LB)
6)	Head circumference at birth (HCB)
7)	Breast circumference at birth (BCB)
8)	Apgar score (APGAR)
9)	Hyper bilirubinemia (BILNEMIA)
10)	Other neonatal abnormalities (OTHABN)
11)	Weight at 3 months (W3M)
12)	Length at 3 months (L3M)
13)	Head circumference at 3 months (HC3M)
14)	Breast circumference at 3 months (BC3M)
15)	History of diseases to 3 months (DIS3M)
16)	Head control (HEADCTRL)
17)	Head lag on pull to sit (HEADLAG)
18)	Stiff extremities (STIFFEXT)
19)	Other abnormalities at 3 months (OTHAB3M)
20)	Weight at 6 months (W6M)
21)	Length at 6 months (L6M)
22)	Head circumference at 6 months (HC6M)
23)	Breast circumference at 6 months (BC6M)
24)	History of diseases to 6 months (DIS6M)
25)	Sitting without support (SITTING)
26)	Rolling over (ROLLING)
27)	Playing with one's hands (HANDPLAY)
28)	Other abnormalities at 6 months (OTHAB6M)

Note: Abbreviation used in text and tables are given in parentheses.

able to perceive during one year after birth (the shortened form of which can be seen in Table 2) and they were collected at the 6-month-checkup.

Prediction systems studied: Three prediction systems of development were studied: 1) a discrimination system simply based on the cumulative number of variables for abnormal states of infants, 2) a system applying a multiple-discriminant analysis, i.e., the method of Hayashi's quantification scaling type 2 (Hayashi, 1954), and 3) a sort of expert system applying the Dempster-Shafer's theory (Shafer, 1976).

In examining the practical application of each system mentioned above, the cross-validity of the criteria used for discriminating between the delayed and the non-delayed infants was simultaneously studied. The procedures were as follows: The 210 subjects are divided into two samples. One, the criterion sample, is used to prescribe the identification rule discriminating between delayed and non-delayed

Table 2. Frequency distribution (%) of infants placed in abnormal category (given in parentheses) of each variable

Variables			Groups: Delayed (N=53)	Groups: Non-delayed (N=157)
(a) *Variables based on developmental records*				
1)	EGA	(<37 wks)	3.8%	5.1%
2)	TYPEDEL	(not spontaneous)	18.9	17.6
3)	ABNORML	(yes)	11.3	7.6
4)	WB	(<10th centile)	17.0	16.6
5)	LB	(<10th centile)	11.3	9.6
6)	HCB	(<10th centile)	17.0	12.7
7)	BCB	(<10th centile)	15.1	13.4
8)	APGAR	(<7)	9.4	2.5
9)	BILNEMIA	(yes)	22.6†	12.1
10)	OTHABN	(yes)	17.0*	5.7
11)	W3M	(<10th centile)	18.9*	7.0
12)	L3M	(<10th centile)	20.8	14.6
13)	HC3M	(<10th centile)	11.3	7.0
14)	BC3M	(<10th centile)	26.4	15.3
15)	DIS3M	(yes)	52.8**	32.5
16)	HEADCTRL	(weak)	26.4**	7.6
17)	HEADLAG	(yes)	15.1*	6.4
18)	STIFFEXT	(yes)	11.3†	4.5
19)	OTHAB3M	(yes)	30.2	35.6
20)	W6M	(<10th centile)	15.1**	3.8
21)	L6M	(<10th centile)	17.0	8.9
22)	HC6M	(<10th centile)	18.9*	8.9
23)	BC6M	(<10th centile)	28.3*	14.6
24)	DIS6M	(yes)	52.8	42.0
25)	SITTING	(no)	37.7*	22.3
26)	ROLLING	(no)	37.7*	22.9
27)	HANDPLAY	(no)	11.3*	3.8
28)	OTHAB6M	(yes)	28.3	24.8
(b) *Variables based on maternal questionnaire*				
1)	Sleeping all day long		1.9	1.3
2)	Waking up at night		13.2	10.8
3)	Sleeping little in daytime		20.8	9.6
4)	Irregular length of sleep		18.9	9.6
5)	Crying little		9.4*	2.5
6)	Crying always		1.9	1.3
7)	Crying loud at night		3.8	1.9
8)	Crying in strange voice		3.8	1.3

Table 2. continued

Variables	Groups: Delayed (N=53)	Groups: Non-delayed (N=157)
9) Sucking weakly	0.0	2.5
10) Taking milk too little	18.9**	5.7
11) Vomiting frequently	9.4	3.8
12) No wish for milk	3.8	1.9
13) Not satisfied after sucking	5.7	2.5
14) No response to mother's call	1.9	0.0
15) Vacant look	3.8†	0.0
16) Not reaching for anything	11.3**	0.0
17) Slow eye movement	1.9	0.6
18) Not looking around curiously	5.7*	0.0
19) Not following visually	5.7*	0.0
20) Not gazing at anything	1.9	0.0
21) No recognition when held up	3.8	3.2
22) Scarcely smiling	1.9	0.0
23) Hardly making voice	1.9	0.0
24) Flabby when held	0.0	0.0
25) Stiff when held	9.4†	3.2
26) Dislike for being held	1.9	0.0
27) Gaining little weight	7.5	3.2
28) Easily falling ill	13.2	5.1
29) Growing up slowly	13.2**	1.3
30) Slow limb movement	3.8†	0.0
31) Convulsive fit	0.0	0.6
32) Too quiet	9.4*	1.3
33) Too nervous	17.0	11.5
34) Strange look in the eyes	1.9	0.0
35) Always in motion	43.4	32.5

Note: In the variables based on maternal questionnaire, the frequencies of affirmative responses are shown.
† $p<.10$; * $p<.05$; ** $p<.01$.

infants, while the other, the validation sample, is used to check the cross-validity of the rule concerned. The two thirds of the subjects randomly drawn from each group construct the criterion sample consisting of 140 infants, 35 of them chosen from the delayed group and 105 from the non-delayed one. On the other hand, the remainders construct the validation sample which consists of 70 infants, 18 of them chosen from the delayed group and 52 from the non-delayed one.

Results

Group comparisons of variables rated as abnormal

Table 2 shows the frequency distribution of infants placed in abnormal categories which were defined (a) by the criteria given in the parentheses in the variables from the developmental checkup records and defined (b) by affirmative responses for the inquiries in the variables from the maternal questionnaire, respectively. As seen in this table, in all except for four variables, the frequencies of the infants placed in abnormal categories were higher in the delayed group than in the non-delayed group. Fisher's exact probability tests were used to assess group differences on frequencies of each variable. Statistically significant or marginally significant group differences were found in about two thirds of the variables of the checkup records and in about one third of those of the questionnaire. This table also indicates that the differences between the two groups become more distinct in the later months.

Table 3. Mean number of variables rated as abnormal for two groups

Group	N	Mean±SD	(Mode)
Delayed	53	8.8±5.8†	(5)
Non-delayed	157	5.0±3.5	(2)

† Significantly larger than non-delayed group at $p<0.001$.

Table 3 presents the mean numbers of variables rated as abnormal for the two groups. The mean number was larger in the delayed group than in the non-delayed group. The difference between two groups was statistically significant ($t(65.5)=4.44$, $p<.001$; calculated by Welch's method, 2-tailed).

Comparisons of 3 systems predicting development

1) A system based on the cumulative number of variables

The results based on the criterion sample revealed that when the cumulative number of variables was used to predict groups, the optimal cutting point for discriminating between the two groups was seven. Thus, the discrimination rule established was that infants having six or less variables rated as abnormal are classified as non-delayed group, whereas infants having seven or more of them are classified as delayed group.

Table 4 gives the classification results of the criterion and the validation samples for the discrimination rule mentioned above. Together with the values of overall agreement, the values of kappa coefficient are also shown in this table as measures indicating validity of discrimination. The left side of the table gives the values for the criterion sample and the right for the validation sample.

Table 4. Classification results obtained from the system based on the cumulative number of valiables rated as abnormal

Actual group	Criterion sample			Validation sample		
	N	Predicted group		N	Predicted group	
		Delayed	Non-delayed		Delayed	Non-delayed
Delayed	35	19 (54.3%)	16 (45.7%)	18	11 (61.1%)	7 (38.9%)
Non-delayed	105	27 (25.7%)	78 (74.3%)	52	13 (25.0%)	39 (75.0%)
% correctly identified:		69.3%			71.4%	
kappa coefficient:		0.26			0.33	

In the criterion sample, though the value of specificity (percent of non-delayed group members correctly identified as non-delayed) exceeds 70%, the value of sensitivity (percent of delayed group members correctly identified as delayed) does not reach 60%, resulting in overall agreement at 69.3% and kappa coefficient at 0.26. These figures seem to indicate ineffective discriminability of the present system.

The cross-validation study based on the validation sample produced a slightly better but substantially similar results, indicating overall agreement at 71.4% and kappa coefficient at 0.33.

2) A system applying a multiple-discrimination analysis

The left side of Table 5 presents the classification results of the criterion sample based on the method of Hayashi's quantification scaling type 2 (Hayashi, 1954), while the right side gives the classification results of the validation sample applying the discrimination rules that came from the preceding analysis on the criterion sample.

As seen on the left side of Table 5, the values of sensitivity and specificity obtained from the criterion sample were very high; 77.1% and 96.2%, respectively,

Table 5. Classification results obtained from the system based on a multiple-discriminant analysis

Actual group	Criterion sample			Validation sample		
	N	Predicted group		N	Predicted group	
		Delayed	Non-delayed		Delayed	Non-delayed
Delayed	35	27 (77.1%)	8 (22.9%)	18	7 (38.9%)	11 (61.1%)
Non-delayed	105	4 (3.8%)	101 (96.2%)	52	14 (26.9%)	38 (73.1%)
% correctly identified:		91.4%			64.3%	
kappa coefficient:		0.76			0.11	

Table 6. Classification results obtained from the system based on Dempster-Shafer's theory

Actual group	Criterion sample			Validation sample		
	N	Predicted group		N	Predicted group	
		Delayed	Non-delayed		Delayed	Non-delayed
Delayed	35	27 (77.1%)	8 (22.9%)	18	14 (77.8%)	4 (22.2%)
Non-delayed	105	17 (16.2%)	88 (83.8%)	52	12 (23.1%)	40 (76.9%)
% correctly identified:		82.1%			77.1%	
kappa coefficient:		0.56			0.48	

thus leading to excellent values of overall agreement at 91.4% and of kappa coefficient at 0.76.

However, when the validation sample was studied by the discrimination rule concerned, the results obtained revealed that the present discrimination rule is of little cross-validity because the measures of discriminability showed such low values as sensitivity being at 38.9%, specificity at 73.1%, overall agreement at 64.3% and kappa coefficient at 0.11. This suggests that the present discrimination system is of little practical applicability.

3) An expert system applying the Dempster-Shafer's theory

The expert system consists of knowledge base and inference machine. The knowledge for predicting developmental delays is stored in the knowledge base as matrices of probability data, representing the relation between predictors and developmental delays. These probability data were derived mainly from our preceding studies (e.g., Adachi & Murai, 1990; Murai et al., 1986). The inference mechanism for processing information of predictors applies Dempster's rule of combination (Dempster, 1967).

The left side of Table 6 shows the classification results obtained from the criterion sample. The values of sensitivity and specificity were very high; 77.1% and 83.8%, respectively. These values were comparable to those obtained by the preceding system based on a multiple-discriminant analysis, in their high values of overall agreement at 82.1% and kappa coefficient at 0.56.

The results based on the validation sample are shown on the right side of Table 6. Though a slight decrease in the value of specificity was observed, the cross-validation study produced substantially similar results to those obtained from the criterion sample, resulting in overall agreement at 77.1% and kappa coefficient at 0.48. These findings suggest that the discrimination rule concerned has satisfactory cross-validity and therefore the present discrimination system is applicable to general practical use.

Discussion

The results confirm that the expert system applying the Dempster-Shafer's theory is the most effective among the three systems of predicting the developmental delays on the basis of early predictors during the first six months. This superiority of the expert system can be explained in at least two ways.

One is that Dempster-Shafer's theory, which enables us to estimate ignorances of information and to integrate them (Shafer, 1976), proves efficient in processing information of predictors. As has been pointed out in other investigations (Broman, 1979; Bronstein & Sigman, 1986; Sigman & Parmelee, 1979), medical, psychosocial, and developmental risk events in early infancy have low predictability for later developmental status of an infant. And as Adachi and Murai (1990) have suggested, professionals have taken various sides of risk events into consideration when interpreting the developmental problem of an infant. These findings mean that early risk events have some uncertainties or ambiguities in the prediction of developmental delays. Systems based on the cumulative number of predictors or applying a multiple-discrimination analysis can't rationally treat ignorances such as uncertainties or ambiguities.

Another explanation of the superiority of the expert system is that the system includes expert knowledge about early predictors in the knowledge base. As mentioned above, predictors have some ignorances in themselves. And we can't always obtain all data which are supposed to be necessary to predict later developmental delays. But knowledge of professionals enables us to deal effectively with erroneous or incomplete data (Hayes-Roth, Waterman, & Lenat, 1983). This is the reason why the system applying a multiple-discrimination analysis didn't have the cross-validity of its discrimination rule.

Thus early prediction systems for developmental delays must treat data including some ignorances. For enhancing the prediction systems to general practical use, we will need to introduce more advanced techniques of knowledge engineering.

References

Adachi, T., & Murai, N. **1990** An analysis of the knowledge of public health nurses on concerning infants' developmental delays. *Tohoku Psychologica Folia.* **49**, 25-32.

Broman, S.H. **1979** Prenatal anoxia and cognitive development in early childhood. In T. Field et al. (eds.), *Infants born at risk.* New York: Spectrum Publications, pp. 29-52.

Bronstein, M.H., & Sigman, M.D. **1986** Continuity in mental development from infancy. *Child Development.* **57**, 251-274.

Dempster, A.P. **1967** Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics.* **38**, 328-339.

Field, T., Hallock, N., Ting, G., Dempsey, J., Dabiri, C., & Shuman, H.H. **1978** A first-year follow-up of high-risk infants. *Child Development,* **49**, 119-131.

Hayashi, C. **1954** Multidimensional quantification-with the applications to analysis of social phenomena-. *Annals of Institute of Statistical Mathematics.* **2**, 121-143.

Hayes-Roth, F., Waterman, D.A., & Lenat, D.B. **1983** An overview of expert systems. In F. Hayers-Roth, D.A. Waterman, & D.B. Lenat (eds.), *Building expert systems*. Massachusetts: Addison-Wesley Publishing Company, pp. 3-30.

Murai, N., Adachi, T., Nihei, Y., Shigemasu, K., Arai, Y., & Murai, N. **1990** Developing a questionnaire asking for maternal information in order to detect infant's developmental disorders. *Japanese Journal of Clinical Pediatrics.* **43**, 713-720 (in Japanese).

Murai, N., Nihei, Y., Adachi, T., Wagatsuma, M., Takahashi, K., Hayasaka, S., Murai, N., Fukuda, T., & Hashimoto, M. **1986** A questionnaire for early detection of developmental disorders utilizing maternal impressions: A retrospective study. *Tohoku Psychologica Folia.* **45**, 114-123.

Parmelee, A.H. Jr., Kopp, C.B., & Sigman, M. **1976** Selection of developmental assessment techniques for infants at risk, *Merrill-Palmer Quarterly.* **22**, 177-199.

Shafer, G. **1976** *A mathematical theory of evidence.* New Jersey: Princeton University Press.

Sigman, M., & Parmelee, A.H. Jr. **1979** Longitudinal evaluation of the preterm infant, In T. Field et al. (eds.), *Infants born at risk.* New York: Spectrum Publications, pp, 193-217.

(Received October 7, 1991)
(Accepted November 18, 1991)