Predicting Post-treatment Scores on the Boston Diagnostic Aphasia Examination

Ann A. VanDemark
University of Michigan, Ann Arbor, Michigan

Several approaches have been taken to the prediction of change in the language behavior of aphasic adults. Darley (1972, 1975) suggested that recovery of language is influenced by a composite of medical, behavioral and environmental variables. For example, younger patients who are in good health and who have had a single instance of brain injury of recent onset are most likely to show significant recovery. Kertesz (1977) identified patterns of recovery associated with different categories of aphasia, and predicted that individuals with Broca's Aphasia would show the greatest magnitude of change. Schuell (1965) and Keenan and Brassell (1974) constructed behavioral profiles based on speech and language performance, and related anticipated changes in the severity of aphasia to early-post-onset profiles. A number of authors (Aten and Lyon, 1978; Wertz, Deal and Deal, 1980, 1981; Porch, Collins, Wertz and Friden, 1980) have discussed predictive methods derived specifically from scores on the Porch Index of Communicative Abilities (Porch, 1971). Matthews and LaPointe (1981) described trend estimations as a method of predicting performance during treatment for aphasia.

Wertz *et al.* (1980) suggested that the most accurate predictions of change in severity of aphasia will probably take into account specific speech and language behaviors as well as medical and biographical data and the effects of treatment. Porch *et al.* also point out that the predictive power of additional variables needs to be determined.

The present study reports the application of multiple regression analysis techniques to the prediction of change in the language performance of aphasic adults. For the purpose of this study, change was defined very narrowly as the difference between scores on the Boston Diagnostic Aphasia Examination administered prior to treatment and the same test administered again at the completion of treatment.

METHOD

The case files of the University of Michigan's Residential Aphasia Clinic (RAC) were searched for records which met the following criteria: the client was an adult (18 years of age or older) who had received a minimum of five weeks of intensive, residential treatment in the RAC during the period from 1976-1981. Medical reports in the client's file indicated a single instance of left hemisphere injury occurring not less than six months before admission to the RAC. A complete record for the Boston Diagnostic Aphasia Examination (Goodglass and Kaplan, 1974) administered both before and after RAC treatment, was present in the client's file. Seventy-five client records met the criteria. The average age of the clients was 41 years 6 months with a range from 19 to 76 years. There were 21 females and 54 males.

The following data were abstracted from each record:

BDAE-1 The total score on the Boston Diagnostic Aphasia Examination administered prior to treatment in the RAC.

BDAE-2 The total score on the Boston Diagnostic Aphasia Examination administered at the completion of treatment in the RAC.

Individual subtest scores from BDAE-1 and BDAE-2.

Medical and biographical information as follows: 1) Age at onset of aphasia, 2) age at start of treatment in the RAC, 3) months post onset of aphasia at start of treatment in the RAC, 4) weeks of treatment at another facility prior to RAC enrollment, 5) months between administration of BDAE-1 and BDAE-2, 6) duration of treatment in the RAC, 7) years of education, 8) sex.

The data were submitted to stepwise, forward-selection multiple regression analysis, using a program available in the Michigan Interactive Data Analysis System (MIDAS) (Fox and Guire, 1976). Multiple regression uses the values of two or more independent variables to predict the value of a dependent variable which contributes most to the variance of the dependent variable and calculates regression statistics. It repeats this process until all of the variables that contribute significantly to the variance of the dependent variable have been selected. The significance level is selected by the investigator prior to the analysis; for the present investigation the .01 level was chosen.

Statistics available from multiple regression analysis include:

1. The multiple correlation coefficient $\underline{R}$ which represents the magnitude of the relationship between the dependent variable and the composite of the independent variables.

2. The coefficient of determination $\underline{R}^2$ which is an estimate of the proportion of the variance of the dependent variable which is accounted for by the composite of the independent variables.

3. The partial regression coefficient $\underline{r}$ which estimates the relative contribution of each independent variable to the variance of the dependent variable.

RESULTS

Seventy-one of the 75 clients included in the sample showed positive change in test scores from BDAE-1 to BDAE-2. The average increase was 61 points, which was significant beyond .001. Test score data are summarized in Table 1.

Table 1. Pre- and post-treatment scores on the Boston Diagnostic Aphasia Examination for 75 aphasic adults.

	Mean	Range	Standard Deviation
Pre-Treatment (BDAE-1)	305.93	81-500	115.4
Post-Treatment (BDAE-2)	361.69	119-513	103.86
(BDAE-2) - (BDAE-1)	60.69	3-181	42.35

Initially, four multiple regression analyses were carried out. They are summarized in Table 2. It will become apparent that these four analyses reflect an element of trial-and-error in the search for a predictive equation, fostered by the ease and speed with which a computer can carry out complex manipulations of data. All four analyses are reported here despite their redundancy because their results and the process they represent are instructive.

In the first multiple regression analysis, the variables which contributed significantly to the prediction of BDAE-2 were: BDAE-1, age at enrollment in RAC, and months between BDAE-1 and BDAE-2. When all three variables were included, the multiple regression coefficient was $R = .94$ and the coefficient of determination was $R^2 = .88$. It appeared that three broad categories of information were predictive: language performance, age, and time.

In the second multiple regression analysis, the total BDAE-1 was eliminated from the list of independent variables and replaced by individual BDAE-1 subtest scores. Three language scores contributed significantly to the prediction of BDAE-2: confrontation naming, body-part identification, and spelling to dictation. Age and time remained as significant contributors. The multiple regression coefficient was $R = .95$ and the coefficient of determination was $R^2 = .90$.

An important question to be addressed was whether data available <u>at the time of an initial evaluation</u> would permit prediction of the score on the post-treatment BDAE. Since information about the number of months that will elapse before the post-treatment evaluation takes place and the length of time an individual may participate in the RAC are not available at the time of an initial evaluation, two multiple regression analyses were carried out eliminating the variables related to time. In multiple regression #3, no other variable replaced "months between BDAE-1 and BDAE-2. The correlation coefficients were the same as those reported at Step 2 in multiple regression #1. In the fourth multiple regression analysis, spelling to dictation was replaced by automatic sequencing, reflecting a minor adjustment in the significance levels for those two variables. The time variable was not replaced by any other. The multiple regression coefficient was $R = .94$ and the coefficient of determinations was $R^2 = .88$.

In each of these four analyses, age at the beginning of treatment in the RAC was a significant contributor to the variance of the dependent variable, its negative correlation indicating that younger patients are likely to achieve higher post-treatment scores. Time, as reflected by months between BDAE-1 and BDAE-2 was a significant factor; when it was removed from the array of independent variables, no other replaced it. Individual language subtest scores were at least as effective and probably more informative than BDAE-1 total test scores.

Data from multiple regression analysis can be used to derive an equation which will predict the individual values of the dependent variable. The data from multiple regression #4 were entered into the prediction formula:

$$Y' = a + b_1X_1 \ . \ . \ . + b_jX_j \quad \text{where}$$

$$b = \frac{s_y}{s_x} \qquad \text{and} \quad a = \overline{Y} - b_1\overline{X}_1 - \ . \ . \ . \ . - b_jX_j$$

as described by Kerlinger (1973). By entering client scores on the four

Table 2. Summaries of four multiple regression analyses.

Multiple regression #1. Dependent variable: BDAE-2. Independent variables: BDAE-1 total score, medical and biographical data.

Step	Variable Entered	Beta Weight	r	R	R^2
1.	BDAE-1	.90	.92	.92	.85
2.	Age enrolled in RAC	-.14	-.37	.93	.87
3.	Months between BDAE-1 and BDAE-2	.11	.30	.94	.88

Multiple regression #2. Dependent variable: BDAE-2. Independent variables: BDAE-1 subtest scores, medical and biographical data.

Step	Variable Entered	Beta Weight	r	R	R^2
1.	Confrontation naming	.65	.89	.90	.80
2.	Body Part Identification	.24	.44	.92	.84
3.	Months between BDAE-1 and BDAE-2	.17	.39	.93	.87
4.	Age enrolled in RAC	-.18	-.37	.94	.89
5.	Spelling to Dictation	.15	.33	.95	.90

Multiple regression #3. Dependent variable: BDAE-2. Independent variables: BDAE-1 total score, medical and biographical data <u>excluding</u> months between BDAE-1 and BDAE-2 and duration of treatment in RAC.

Step	Variable Entered	Beta Weight	r	R	R^2
1.	BDAE-1	.88	.91	.92	.85
2.	Age enrolled in RAC	-.14	-.36	.93	.87

Multiple regression #4. Dependent variable: BDAE-2. Independent variables: BDAE-1 subtest scores, medical and biographical data <u>excluding</u> months between BDAE-1 and BDAE-2 and duration of treatment in RAC.

Step	Variable Entered	Beta Weight	r	R	R^2
1.	Confrontation Naming	.89	.90	.90	.80
2.	Body Park Identification	.24	.45	.92	.84
3.	Age enrolled in RAC	-.15	-.35	.93	.86
4.	Automatic Sequencing	.13	.31	.94	.88

variables identified in multiple regression #4 into the equation, an individual's performance on BDAE-2 can be predicted. The accuracy of individual prediction is, of course, far more important and informative than the values of $\underline{R}$ and $\underline{R}^2$. The predicted BDAE-2 scores were calculated for the clients in the study population. Twenty (27%) of the predicted scores were within ± 10 points of the actual scores, and 37 (49%) were within ± 20 points of the actual score. The differences between predicted and actual scores were as small as -0.06 and as large as 149.95. In the latter case, the predicted BDAE-2 was 230.44 and the actual BDAE-2 was 380. The client's post-treatment performance far surpassed his predicted score. There was a tendency toward small over-estimates and larger, more variable under-estimates of BDAE-2. The actual prediction equation for the data from multiple regression #4 are presented in Table 3.

Table 3. Prediction equation for the data from multiple regression #4.

Y' = 204.75 + 1.68 (confrontation naming score) + 5.62 (body part identification score) + 6.75 (automatic sequencing score) - 1.034 (age at enrollment in RAC).

Two subsets of data were examined in more detail; The data for those subjects (n=20) whose predicted BDAE-2 score was within ± 10 points of their actual BDAE-2 score, and the data for those subjects (n=14) whose predicted BDAE-2 score exceeded or fell short of the actual score by 40 points or more. Scores for both groups on BDAE-1 and BDAE-2 and the four variables included in the predictive equation are summarized in Table 4. As a group, the subjects for whom prediction of BDAE-2 was reasonably accurate performed significantly better on all of the language measures and were significantly younger than the group for whom prediction of BDAE-2 was not accurate. However, the overlap between the distributions for the variables does not permit the conclusion that prediction is consistently more accurate for less-impaired, younger subjects.

Thus, the variables identified in multiple regression #4 (confrontation naming, body part identification, automatic sequencing, and age at enrollment in the RAC) were highly predictive of post-treatment test scores for some subjects, but not for all.

The next phase of the investigation involved an attempt to identify additional variables which might improve the accuracy of the predictive equation. Client files were reviewed again, and information related to site and nature of brain injury was coded and added to the data set. A fifth multiple regression analysis was performed and is summarized in Table 5.

Comparison of multiple regression analyses #4 and #5 indicates that adding coded data on site and nature of brain injury had no significant effect on the accuracy of prediction of BDAE-2. Word-picture matching was substituted for automatic sequencing at step 4 — apparently the last variable selected for the regression is subject to considerable variation. Examination of the significance levels for the variables not included in the regression equation indicated that spelling to dictation, automatic sequencing, and word-picture matching all had significance levels very close to the cut-off point for inclusion in the equation, so that small changes in the values for other variables tended to move these three in or out of the equation.

Table 4. Scores on the Boston Diagnostic Aphasia Examination and four predictor variables for aphasic subjects whose post-treatment scores were predicted accurately and inaccurately.

	BDAE-1	BDAE-2	Confrontation Naming	Body Part Identification	Automatic Sequences	Age at Enrollment RAC
			Accurate Prediction			
Mean	331.15	389.68	63.47	14.21	6.52	33.10
Median	331.00	406.00	73.00	16.00	8.00	29.00
Range	90-500	235-513	2-105	2-20	1-10	19-57
s.d.	125.54	95.28	35.56	5.49	2.79	12.26
			Inaccurate Prediction			
Mean	219.42	301.64	30.28	9.64	3.42	46.00
Median	222.00	331.50	25.00	11.50	3.50	49.00
Range	81-384	119-462	0-84	0-18	0-8	29-62
s.d.	99.62	127.43	24.41	5.82	2.69	10.74
P	.05	.05	.01	.05	.001	.01

Table 5. Multiple regression #5. Dependent variable: BDAE-2. Independent variables: BDAE-1 subtest scores, medical and biographical data including site and nature of brain injury.

Step	Variable Entered	Beta Weight	r	R	R^2
1.	Confrontation Naming	.61	.87	.87	.77
2.	Body Part Identification	.27	.45	.90	.81
3.	Age Enrolled in RAC	-.15	-.36	.92	.84
4.	Word-Picture Matching	.13	.31	.93	.86

Finally, separate multiple regression analyses were computed for those subjects who had experienced vascular lesions (n = 57) and those who had experienced traumatic head injuries (n = 13). These analyses are summarized in Tables 6 and 7. They contribute significantly to the interpretation of the earlier analyses. For clients with aphasia secondary to vascular lesions, three language subtests predict the post-treatment scores on the BDAE: confrontation naming, complex ideational material, and body part identification. It seems likely that the language dimensions reflected here are word retrieval and auditory comprehension. For a reason that will become clearer shortly, age is no longer a predictor. None of the other biographical or medical variables contributed significantly. The accuracy of prediction is only slightly lower than for the larger group ($R = .91$, $R^2 = .83$).

Table 6. Multiple regression #6. Clients with vascular lesions (n=57). Dependent variable: BDAE-2. Independent variables: BDAE-1 subtest scores, medical and biographical data.

Step	Variable Entered	Beta Weight	r	R	R^2
1.	Confrontation Naming	.56	.69	.86	.74
2.	Complex Materials	.25	.39	.89	.81
3.	Body Part Identification	.21	.35	.91	.83

Multiple regression #7 demonstrates rather astonishingly that a group so small (n=13) can still be highly predictable. Ninety percent of the variance of BDAE-2 is accounted for by the confrontation naming score on BDAE-1. For clients with aphasia secondary to closed head injury, the ability to name to confrontation is strongly predictive of their performance on BDAE-2.

Table 7. Multiple regression #7. Clients with traumatic head injury (n=13). Dependent variable: BDAE-2. Independent variables: BDAE-1 subtest scores, medical and biographical data.

Step	Variable Entered	Beta Weight	r	R	R^2
1.	Confrontation Naming	.95	.95	.95	.90

The variable "age at enrollment in RAC" is conspicuously absent from both analyses, after being conspicuously present in all those preceeding. The reason becomes apparent when the two groups (vascular and traumatic) are compared on the basis of age. The subjects with aphasia secondary to traumatic head injury were significantly younger than those with aphasia secondary to vascular disorder ($\bar{X}_{trauma}$ = 30 years 6 months, $\bar{X}_{vascular}$ = 42 years 2 months, p = .0028). Age is a predictor only when we attempt to predict for both groups of subjects in the same regression analysis.

DISCUSSION

At this stage of the investigation, with at least as many new questions raised as old ones answered, the following conclusions seem appropriate:

(1) Performance by an aphasic client on a post-treatment test of language (BDAE) can be predicted with impressive accuracy, and with frightening inaccuracy.

(2) The single variable which stands out as a predictor is confrontation naming. It is present in every analysis where individual language performances are considered. Variables representing auditory comprehension are also consistently represented for all subjects except those with aphasia secondary to traumatic head injury.

(3) Age is predictive when traumatic head injury and vascular patients are included in the same analysis, but not when they are separated.

(4) It is possible that a combination of variables exists which will predict more accurately than the combination identified in these analyses. It is likely that difficult-to-quantify measures such as motivation and quality of treatment have a role in the prediction of change. It is equally clear that BDAE-2 minus BDAE-1 is not a complete descriptor of change in the language behavior of aphasic clients. Our efforts thus far are merely a beginning.

(5) Statistical prediction of post-treatment test scores has a place in our diagnostic repertoire. We can and should use it — we cannot and must not rely solely upon such number manipulations in making decisions about people with communication disorders.

REFERENCES

Aten, J.L., Lyon, J.G. Measures of PICA subtest variance: A preliminary assessment of their value as predictors of language recovery in aphasi patients. In R.H. Brookshire (Ed.), Clinical Aphasiology Conference Proceedings. Minneapolis, MN: BRK Publishers, 1978.

Darley, F.L. The efficacy of language rehabilitation in aphasia. Journal of Speech and Hearing Disorders, 37, 3-21, 1972.

Darley, F.L. Treatment of acquired aphasia. In W.J. Friedlander (Ed.), Advances in Neurology, Vol. 7: Current Reviews of Higher Nervous System Deficiencies. New York: Raven Press, 111-145, 1975.

Fox, D.J. and Guire, K.E. Documentation for MIDAS (3rd Ed.) Statistical Research Laboratory, University of Michigan, 1976.

Goodglass, H. and Kaplan, E. Assessment of Aphasia and Related Disorders. Philadelphia: Lea and Febiger, 1972.

Keenan, J.S. and Brassell, E.G. A study of factors related to prognosis for individual aphasic patients. Journal of Speech and Hearing Disorders, 39, 257-269, 1974.

Kerlinger, F.N. Foundations of Behavioral Research (2nd Ed.) New York: Holt, Reinhart and Winston, Inc., 1973.

Kertesz, A.S. and McCabe, P. Recovery patterns and prognosis in aphasia. Brain, 100, 1-18, 1977.

Matthews, B.A.J. and LaPointe, L.L. Determining rate of change and predict ing performance levels in aphasia therapy. In R.H. Brookshire (Ed.), Clinical Aphasiology Conference Proceedings. Minneapolis, MN: BRK Publishers, 1981.

Porch, B.E. The Porch Index of Communicative Ability. Palo Alto: Consulting Psychologists Press, 1967.

Porch, B.E., Collins, M., Wertz, R.T., and Friden, T.P. Statistical prediction of change in aphasia. Journal of Speech and Hearing Research, 23, 312-321, 1980.

Schuell, H. Differential Diagnosis of Aphasia with the Minnesota Test. Minneapolis, MN: University of Minnesota, 1965.

Wertz, R.T., Deal, L., and Deal, J. Prognosis in aphasia: investigation of the high-overall and the short-direct methods to predict change in PICA performance. In R.H. Brookshire (Ed.), Clinical Aphasiology Conference Proceedings. Minneapolis, MN: BRK Publishers, 1980.

Wertz, R.T., Deal, L.M., and Deal, J.L. Clinical significance of the PICA high-low gap. In R.H. Brookshire (Ed.), Clinical Aphasiology Conferen Proceedings. Minneapolis, MN: BRK Publishers, 1981.

DISCUSSION

Q: Do you know of any psychological measure that would tap into an aphasic's motivation?

A: I'm sure there are some, but I expect they are paper and pencil tests which would be difficult to use with an aphasic population.

Q: I'm interested in the mean age of your stroke population. It seems to me that your mean age is lower than the typical age for stroke patients. I'm wondering if you would like to speculate a little bit on the possibility that you had some patients in the stroke group who were somewhat atypical of the aphasic population.

A: This is a very typical group for our clinic, but our clinic is rather selective about the clients that come in—for example they have to be independent in daily living skills. Yes, I think it is a younger population, even though they are vascular patients.

Q: Was size of lesion ever a potential variable in your analysis?

A: No, we didn't have that kind of data available.

Q: Can you describe the type or severity of aphasia of those aphasic persons for whom you were able to predict well?

A: They were younger and they did tend to perform better on the four variables in the predictive equation, and on the other language variables as well.

Q: Did the months between BDAE-1 and BDAE-2 reflect longer treatment stays for those patients?

A: It reflected two things. There are varying lengths of time between an initial evaluation and the actual beginning of therapy in the aphasia clinic. Some clients receive therapy in their home community in the interim, some do not, and the frequency of therapy varies considerably. So the measure reflects the length of time between application and admission as well as the time actually spent in the aphasia clinic.

Q: Were these all stable aphasic patients?

A: Yes. They were at least six months post onset.

Q: Does the increase in the percent of variance accounted for in the first multiple regression by the last factor—does it really only account for an additional one percent of the variance?

A: All three of those variables contribute to the final prediction.

Q: What are the criteria for determining the length of treatment?

A: Continued progress, continued interest on the part of the patient, continued availability of financial support.

Q: What is a typical stay?

A: Typical stay is 30 weeks. The range represented in this sample was 5-55 weeks, but both extremes of that range were quite unusual.