Regression Equations for Predicting Premorbid Performance on Executive Test Measures by Persons with Traumatic Brain Injuries

Mairead A. Harnett, Hamish P.D. Godfrey, and Robert G. Knight
University of Otago

Regression equations, using as predictors demographic variables and scores on tests resistant to the effects of brain injury, are often used to calculate predicted test scores on tests used in neuropsychological assessment. The discrepancy between a predicted and an obtained score can be used in the diagnosis of cognitive deficits resulting from brain damage. The aim of the present study was to develop regression equations that would allow the prediction of scores on tests sensitive to executive deficits and to apply these to data from persons with brain injuries. A total of 100 persons from the community were tested and the resultant equations were used to predict the scores of 40 persons with traumatic brain injuries. For those equations with multiple correlations in the order of .50, the numbers of persons classified using either the equations or the published normative tables were comparable. Procedures for determining the abnormality of an individual's test score discrepancy described by Crawford and Howell (1998), based on the output from multiple regression procedures, are discussed. Where normative data from a local sample are available, use of regression equations to determine the abnormality of a discrepancy can provide a useful means of validating conclusions based on the application of US or UK norms.

In clinical practice, neuropsychologists often need to determine whether a test score obtained from a person with neurological damage differs from what might have been expected if no brain injury had occurred. One means of achieving this is to estimate a client's expected score with a regression equation, using as predictors demographic variables such as age or socio-economic status, other test scores obtained either concurrently or pre-injury, or some combination of both. The standardised difference between the expected and obtained scores can then be used as an index of acquired cognitive impairment. In effect, regression equations are being used in this situation as an alternative to consulting conventional normative tables (Crawford & Howell, 1998).

Most manuals of neuropsychological tests contain tables of norms, often stratified by age and, in some instances by education levels, that allow evaluation of a client's obtained score in terms of a distribution of scores from a representative sample of healthy individuals. An abnormal score is one that lies at the extreme of the distribution. Although normative tables are a useful method of conveying information about expected scores, there can be advantages to developing regression equations. For example, the size of the sample is not so significant for regression-based norms, provided it is sufficiently large and diverse enough to provide stable correlations between the test and predictors. This is a significant advantage where local norms are being constructed and the resources for collecting a large and representative sample may not be available. For example, in New Zealand it may often be more cost effective to develop regression equations as an aid to test interpretation than to renorm neuro-psychological tests standardised in the United Kingdom or the United States. There are other advantages to using regression equations. The predicted or expected score is based on continuous variables rather than grouped variables, that is, on the client's actual age, for example, rather than on the age band in which they happen to fall. Thus the use of regression equations to estimate test scores can be more precise than referring to tables, and where more than one variable is used to stratify the sample, less cumbersome. As Crawford and Howell (1998) have observed, applying regression equations means that "... an individual's predicted score reflects his/her particular combination of demographic characteristics. Such an approach is in keeping with the emphasis placed on individual versus normative comparison standards in neuropsychological assessment (p. 755)."

There are numerous examples in the literature where regression equations have been constructed to convey normative information for the assessment of individual cases (e.g., Crawford, Allen, Cochrane, & Parker, 1990; Crawford & Howell, 1998; Crawford, Parker, Stewart, Besson, & Delacey, 1989; O'Brien, Godfrey, Freeman, & Perkins, 1999). Two classes of variables have commonly been used to estimate premorbid cognitive ability. One is "hold" tests, that is, measures based on over learned skills such as reading aloud or the application of semantic knowledge, which are relatively resistant to the effects of brain injury in non-aphasic patients. Of such tests, the most widely researched is the National Adult Reading Test (NART; Nelson & Willison, 1991), a measure of word reading ability. Performance on the NART has been shown to be largely unaffected by mild to moderate dementia (O'Carroll, Blaikie, & Whittick, 1987) and closed head injury (Crawford, Parker, & Besson, 1988; Moss & Dowd, 1991). Demographic variables such as age, education level, and pre-injury socio-economic status are the other class of variable used to estimate premorbid functioning in regression equations (e.g., Barona, Reynolds, & Chastian, 1984; Franzen, Burgess, and Smith-Seemiller, 1989; Karzmark, Heaton, Grant, & Mathews, 1985). For example, Crawford and Allen (1997) found that regression equations developed using the demographic variables of education, race, occupation, and age had R^2 values of .53, .53, and .32 for predicting WAIS-R Verbal IQ, Performance IQ, and Full Scale IQ, respectively.

More recent approaches to estimating premorbid cognitive ability employ multiple regression equations to predict specific test scores for an individual using as predictors both demographic variables and hold tests (Crawford, 1992; Crawford, Moore, & Cameron, 1992). Combining demographic variables and performance on hold tests has produced the most accurate estimates of premorbid functioning currently available (Crawford, Stewart, Parker, Besson, & Cochrane, 1989; Crawford et al., 1992; Freeman & Godfrey, 1999; Krull, Scott, & Sherer, 1995; Van den Broek & Bradshaw, 1994), but to date, this procedure has not been widely applied to the assessment of executive deficits. Exceptions to this are the studies by Crawford et al. (1992), in which NART scores were successfully used to predict performance on a measure of verbal fluency, and Crawford, Obonsawin, and Allen (1998), where WAIS-R IQ, NART scores, and age were used to predict premorbid Paced Auditory Serial Addition Task (PASAT) scores.

The primary purpose of the present study was to develop regression equations for predicting scores on tests of neuropsychological functioning and to test their effectiveness on sample of persons with traumatic brain injury (TBI). TBI is a prevalent and expensive health problem in New Zealand (Caradoc-Davies & Dixon, 1995), commonly resulting in difficulties in the initiation and regulation of behaviour, often referred to as dysexecutive syndrome (Baddeley & Wilson, 1988). Impairments in executive functioning may present as difficulties with self-initiation, goal-oriented planning, sequencing and organising behaviour, and monitoring outcomes accurately (e.g., Milner, 1964; Stuss & Benson, 1986; Tate, 1999; Tranel, Anderson, & Benton, 1994). Research findings indicate that dysexecutive syndrome is often associated with loss of social autonomy and the inability to return to work after TBI (Mazaux, Masson, Levin, Alaoui, Maurette, & Barat, 1997). Because of the significant psychosocial impact of dysexecutive syndrome and the need to put in place appropriate rehabilitation strategies as soon as possible, timely and accurate diagnosis is a priority.

As part of the process of documenting the deficits resulting from TBI, clinical neuropsychologists frequently administer standardised tests of executive function. Accurate interpretation of the scores from the tests used in the diagnosis of dysexecutive syndrome is dependent upon the availability of normative data stratified for relevant demographic variables, such as education and age. For many such tests, however, the published norms are limited and often the only demographic variable taken into account is age. Although age-band norms are useful, frequently they do not allow sufficiently precise score interpretation. For example, because of the range of cognitive ability in the general population, a score that lies in the normal range of functioning may represent a substantial decline for a high functioning person but no change in the normal functioning for someone of lesser ability. For example, Borkowski, Benton, and Spreen (1967) showed that on a common measure of executive functioning (verbal fluency), brain damaged patients with higher than average IQs performed better than those patients with lower than average IQs. A further issue for clinicians practising in New Zealand is that the normative data for neuropsychological tests typically come from US or UK samples, and norms based on demographic variables such as education or scores on tests of verbal-educational abilities may not generalise from one country to another.

Although the primary goal of the present study was to develop regression equations to predict premorbid performance on measures of executive functioning, a more general aim was to investigate the usefulness of constructing norms for neuropsychological tests in this manner. Developing regression equations can be a more efficient way of constructing local norms, than preparing detailed age- or education-band normative tables. The general research strategy involves administrating neuropsychological tests to a diverse group of healthy adults in the community and computing correlations between the test scores and relevant premorbid predictors. Assuming that the sample is large and contains a range of individuals of differing ability, these correlations will be relatively stable and valid estimates of the relationships between variables in the general population. The resulting correlation matrix is used to determine a regression model comprising the best predictors for each test score. These regression equations can then be applied to the test scores of persons who have acquired brain injury. In each case, a predicted test score is calculated using the appropriate regression equation, and the standardised discrepancy between obtained and expected scores computed. The magnitude of the discrepancy can then be expressed precisely as a percentile, by consulting tables of areas under the normal curve, or cut-off scores determined for specific percentiles.

In the present study, a group of 100 persons with diverse backgrounds were recruited from the community, and correlations between their executive test scores and relevant test and demographic predictors computed. Age, gender, years of education, and NART scores were selected as relevant predictors on the basis of their widespread use in comparable studies. Data from this group were used to determine the regression equations. A sample of persons with TBI and ongoing neurological problems was then recruited, and the same data collected. As a check on the validity of the neuropsychological tests as measures of the cognitive impairment resulting from brain injury, matched samples of TBI and community participants were compared on the executive tests. This process also served to establish that the persons with TBI had current evidence of neuropsychological deficit. The regression equations based on the results from the community sample were then applied to the data from the TBI group. The number of persons identified as impaired (below the 10th percentile) using these equations and the test norms was determined, as a preliminary examination of the diagnostic sensitivity of the regression procedure.

Method

Participants

Community sample. A total of 100 participants (43 males and 57 females) with an average age of 39.10 years (SD = 11.75, range = 17-64) and no history of traumatic brain injury, neurological, or developmental disorder were recruited from local community organisations (76%), undergraduate students (14%), and siblings or first cousins of the traumatically brain injured (TBI) participants (9%). The participants had an average of 14.80 years (*SD* = 3.26, range = 9-20) of education, and an average NART error score of 13.92 (*SD* = 7.64, range = 1-36). Their average score on the Beck Depression Inventory- II (BDI-II) was 5.9 (*SD* = 6.5). A total of 11% had no formal academic qualifications; the highest qualification for 39% was achieved at school, 37% had a university degree, and 13% a vocational qualification.

TBI sample. Forty persons (23 males and 17 females) with an average age of 42.10 years (*SD* = 11.28), were recruited through the Head Injury Societies and Brain Injury Associations of Canterbury, Otago, Southland, Westland and Nelson, New Zealand. These participants had all suffered a documented TBI, and had no history of other neurological disease, psychiatric condition, or developmental disorder. Causes of injury were classified as motor vehicle accident (66%), assault (11%), fall (9%), blow from a moving object (9%), and other (5%). All persons with TBI were at least 15 months post injury (*M*=110.15 months, *SD* = 82.22), and their average age at the time of injury was 32.48 years (*SD* = 12.60). Hospital records were available for 39 participants and the average length of hospitalisation was 64.72 days (*SD* = 73.33). On average, participants were unconscious for a period of 428.76 hours (*SD* = 82.22), and had recorded periods of post traumatic amnesia averaging 105.06 days (*SD* = 258.92). Most participants had suffered an injury in the severe range, which left them with some degree of neurological impairment and ongoing contact with support groups.

The TBI group had an average of 12.65 years of education (*SD* = 2.99). A total of 23% had no formal academic qualifications; the highest qualification for 62% was achieved at school, 5% had an undergraduate degree, and 8% a vocational qualification. At the time of testing 42% were unemployed; at the time of injury 5% had been unemployed. The average BDI-II score was 14.45 (*SD* = 8.92) and the mean error score on the NART was 18.98 (*SD* = 9.42).

Measures

National Adult Reading Test. The NART (Nelson & Willison, 1991) is a reading test comprising 50 words of irregular pronunciation (e.g. psalm, cellist), designed to estimate premorbid intellectual level. The NART was administered to all participants as a psychometric predictor of the executive test scores.

Beck Depression Inventory – II (BDI – II; Beck & Steer, 1996). Because the prevalence of depression is high in persons with TBI (Godfrey, Partridge, Knight, & Bishara, 1993), the BDI II was administered to investigate the possible concurrent effects of low mood on test performance. This inventory is composed of 21 items measuring symptoms of depression.

Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Tailey, Kay, & Curtis, 1993). A computer-administered version of the WCST (Computer Version 3; Heaton & PAR staff, 1999) was used. Scores on the WCST, particularly rate of perseverative responding, have been demonstrated to be sensitive to the effects of TBI (Cockburn, 1995; Heaton, 1981; Heaton et al., 1993). The comparability of tester- and computer-administered versions has been demonstrated for normal (Artiola i Fortuny & Heaton, 1996) and psychiatric (Hellman, Green, Kern, & Christensen, 1990) groups. The WCST produces a number of scores; for the purposes of this study, number of categories, perseverative responses, perseverative errors, and total errors were used as indicators of test performance.

Trail Making Test (TMT; Reitan, 1955). The TMT is a two-part divided-attention and speed of information processing task. Both Parts A and B of the TMT have been demonstrated to be sensitive to the effects of TBI (Johnstone, Leach, Hickey, Frank, & Rupright, 1995).

Controlled Oral Word Association Task (COWAT; Spreen & Benton, 1969). The COWAT is a measure of verbal fluency sensitive to the effects of TBI (e.g., Cockburn, 1995; Raskin & Rearick, 1996), in which participants are asked to generate as many words beginning with F, A, or S in a limited period of time.

Semantic Fluency (Goodglass & Kaplan, 1983). The Semantic Fluency test requires participants to generate names from a specified category, in this case, animals, for 60 seconds. Semantic fluency has been shown to be sensitive to the effects of TBI (e.g., Raskin & Rearick, 1996).

Short Form of the Booklet Category Test (BCT). Russell and Levy (1987) have constructed a shorter 95-item booklet version of the original Category

Test (Halstead, 1947), which requires the same application of rule learning and shift in conceptual set as the original. They reported a correlation of 0.97 between the two versions. The BCT has demonstrated sensitivity to brain damage (e.g., Mercer, Harrell, Miller, Childs, & Rocker, 1997).

Design Fluency Task (DFT; Jones-Gotman & Milner, 1977). The DFT is a measure of non-verbal fluency requiring the generation of unique designs in a limited period of time. The free condition requires the participant to draw as many different designs as they can. In the fixed condition participants are instructed to draw as many designs as they can, using only four lines. Designs that represent actual objects (e.g. a person or a house) or regular geometric figures with names (e.g., square), and scribbles, are not permitted. The total number of novel designs (total minus unacceptable designs) was recorded for each condition.

Procedure

All testing was conducted in the participants' homes. They first completed the BDI-II and were interviewed to obtain information regarding their current circumstances and neurological history. The participants were then administered the neuropsychological measures in the following order: WCST, NART, Semantic Fluency test, TMT, COWAT, the Design Fluency, and the BCT. The computerised version of WCST was administered using a PC laptop with a 14-inch screen. All statistical analyses were performed using Statistical Package for the Social Sciences (SPSS, version 10).

Results

Comparison between TBI and Community Samples

In order to determine which of the neuropsychological measures were most sensitive to the long-term effects of TBI, data from the two samples were compared. As the primary purpose was of the study was to develop regression equations on a diverse community sample and apply them to a separate sample of persons with TBI, no attempt was made when recruiting participants to match the two groups. It was possible, however, to match each of the 40 TBI participants with a member of the community sample for age at testing, education level, and gender.

Means, standard deviations, and the results of independent sample t-tests comparing the two groups are presented in Table 1. Means and standard deviations for the total community sample are given as part of Table 2. Effect Sizes (ES) for the two matched groups were calculated by dividing the group mean difference by the control group standard deviation and are presented in Table 1.

There were no significant differences in age, education, and NART errors between the two groups, however, as expected the TBI group scored significantly more highly on the BDI-II. There were no significant correlations between BDI-II scores and test scores for the community sample, but for the persons with TBI, there were significant correlations between depression scores and the TMT (Part A), r (38) = .35, $p < .05$, and Semantic Fluency, r (38) = -.37, $p < .05$. The neuropsychological tests that were most sensitive to the effects of TBI were the Semantic Fluency test and TMT Parts A and B, with ES values greater than 1. There was also a significant difference between groups on the COWAT, but no significant differences for the Design Fluency, WCST, and the BCT scores. These findings confirmed the presence of some degree of cognitive impairment in the TBI group.

Multiple Regression Analyses

The construction of regression-based norms requires an estimation of the magnitude of the correlations between the tests and predictor variables in the

Table 1. Comparison of Matched Groups (*n* = 40) of TBI and Community Participants

	TBI		Control			
	M	*SD*	*M*	*SD*	*t*	*ES*
Age	42.10	11.28	41.70	11.67	<1.0	.03
Years Education	12.65	2.99	13.46	2.56	1.31	.10
BDI-II	14.14	8.92	6.60	5.49	4.74**	1.43
NART Errors	19.98	9.42	16.20	8.23	1.40	.34
WCST						
Categories	5.00	1.65	5.31	1.56	<1.0	.20
Perseverative Responses	18.46	14.87	12.95	10.18	1.89	.54
Perseverative Errors	16.59	12.47	11.90	8.63	1.77	.54
Total Errors	31.54	19.67	23.38	17.95	1.89	.45
Semantic Fluency	19.77	5.99	24.60	4.67	4.02**	1.03
TMT Part A (secs)	46.55	17.05	29.40	11.16	4.08**	1.18
TMT Part B (secs)	98.47	42.05	58.65	20.76	5.35**	1.92
BCT Errors	25.28	12.74	23.72	13.48	<1.0	.11
COWAT	34.53	14.45	44.77	12.93	3.32**	.49
Design Fluency						
Free novel total	16.29	10.47	19.15	8.03	1.34	.36
Fixed novel total	14.51	7.92	16.95	7.05	1.45	.35

Note: BDI = Beck Depression Inventory; ES = Effect size; NART = National Adult Reading Test; WCST = Wisconsin Card Sorting Test; TMT = Trail Making Test; COWAT = Controlled Oral Word Association Test; BCT = Booklet Category Test.

* $p < .05$ ** $p < .01$

general population of persons without neurological impairment. To this end the correlations between the relevant variables were determined for the community sample and then submitted to a stepwise multiple regression analysis. Gender, age, years of education, and NART errors were used as predictors of each of the measures listed in Table 1 (except number of WCST categories, which was excluded because it was skewed to an extent that made it unsuitable for entry as a dependent variable in a regression analysis). In Table 2, the correlations between the variables entered in the multiple regression analyses are presented. Generally the correlations were modest, with NART errors and age being the variables most commonly associated with the executive test variables. Table 3 shows the regression equations produced, standard errors, and multiple correlation estimates for the community sample. R^2 values were also relatively modest, with no equation explaining more than 30% of the test score variance.

Table 2. Correlations between the Dependent and Predictor Variables in the Regression Analysis

	M	*SD*	Correlations			
			Gender	Age	Educ	NART
WCST						
Perseverative Errors	10.31	7.83	-.06	.34**	-.20	.33**
Perseverative Responses	11.30	9.929	-.04	.34**	-.21	.30**
Total Errors	19.88	15.89	-.02	.30**	-.21	.37**
Semantic Fluency	26.63	5.76	.04	-.09	.31**	-.33**
TMT Part A (secs)	26.33	9.71	-.27**	.16	-.09	.13
TMT Part B (secs)	51.26	17.93	-.14	.26**	-.31**	.32**
BCT Errors	20.84	12.07	.17	.37**	-.25*	.27
COWAT total	46.88	11.63	.16	.12	.25*	.47**
Design Fluency						
Free Novel Total	20.76	8.85	.01	.01	.29**	-.19
Fixed Novel Total	18.35	7.95	.15	-.20*	.17	-.22
Age	39.10	11.75	.16	-	-.03	-.22*
Educ	14.80	3.26	.03	-.03	-	-.52**
NART	13.92	7.64	-.18	-.22	-.52**	-
Gender	-	-	-	.16	.03	-.18

Note: NART = National Adult Reading Test; WCST = Wisconsin Card Sorting Test; TMT = Trail Making Test; COWAT = Controlled Oral Word Association Test; BCT = Booklet Category Test; Educ = Years Education.

* $p<.05$ ** $p<.01$

Comparison between predicted and obtained scores

As expected, there were no significant differences between mean predicted and obtained scores for the 10 test scores used as dependent variables in the regression analysis for the community sample, paired sample t (97) < 1.0 in each case. For the TBI sample, however, the discrepancies between obtained and predicted test scores were significant for WCST perseverative errors, t (36)=2.35, $p< .05$, Semantic Fluency, t (39) = 6.07, $p< .001$, TMT (Part A), t (39) = 5.26, $p< .001$, TMT (Part B), t (39) = 6.57, $p< .001$, COWAT, t (39) = 4.59, $p< .001$, and Design Fluency (Free total novel), t (38)=3.45, $p< .001$. There were no significant differences between obtained and predicted scores for the remaining test scores.

Since there is no independent criterion of executive dysfunction, it is not possible to test directly the validity of diagnoses based on the regression equations. As an indirect check, however, diagnoses made on the basis of the regression equations were compared with those made using published normative tables, by determining the number of persons falling beyond the 10th percentile using both methods. For the regression equations, z-scores for each participant in both groups were calculated by subtracting the obtained from the predicted scores, and dividing by the standard error of the estimate (these values are given in Table 3). All individuals with z-scores of 1.3 or greater (i.e., at or below the 10th percentile, one-tailed) were classified as impaired. The percentage of TBI and community participants falling below the 10th percentile using published norms (Heaton & PAR staff, 1996; Russell & Levy, 1987; Spreen & Strauss, 1998; Tombaugh, Kozack, & Rees, 1999) was then determined for each measure. The results for both methods, with the cut off scores used to determine the magnitude of a discrepancy at the 10th percentile for the regression equations, are given in Table 4. Percentiles of less than 20 were not available for TMT Part B norms so although the regression equations identified those falling below the 10th, the norms identified those falling below the 20th percentile. Scores at the 10th percentile for the Design Fluency test were estimated from the means and standard deviations presented in Spreen and Strauss (1998). No comparison was made for the BCT, because this test simply classifies individuals as impaired (errors = 23), and percentile scores are not available.

As is apparent in Table 4, the classification rates for both methods were comparable for the community sample, and as would be expected, much lower for this sample than for the TBI group. For the TBI sample, classification rates using the

Table 3: Regression Equations with the Demographics Variables and NART Errors as Predictors (Community Sample, n = 100)

Dependent Variables	Equation	*R*	R^2	SE_{est}	*F*
WCST					
Perseverative Response	-8.25 + .47 NART + .33 Age	.51	.26	8.07	16.63**
Perseverative Errors	-6.54 + .42 NART + .28 Age	.53	.28	6.73	18.07**
Total Errors	-13.47 + .93 NART + .52 Age	.53	.28	13.64	18.15***
Semantic Fluency	30.11 - .25 NART	.33	.11	5.47	12.09***
TMT Part A (secs)	28.82 - 5.83 Gender + .17 Age	.34	.11	9.24	6.15**
TMT Part B (secs)	17.75 + .93 NART + .53 Age	.47	.22	16.03	13.41***
BCT Errors	-5.51 + .46 Age + .58 NART	.51	.26	10.47	16.40***
COWAT_Total	56.76 - .71 NART	.47	.22	10.34	26.63***
Design Fluency					
Free Novel Total	9.01 + .79 Educ	.29	.09	8.50	8.91**
Fixed Novel Total	29.03 - .20 Age - .22 NART	.31	.10	7.61	5.36**

Note: NART = National Adult Reading Test; WCST = Wisconsin Card Sorting Test; TMT = Trail Making Test; BCT = Booklet Category Test; COWAT = Controlled Oral Word Association Test; Educ = Years Education.

regression equations were inexplicably low for the Design Fluency test, and this suggests that the regression equations for measures from this test, which also had low R^2 values, may not be valid. The equations for the WCST, TMT, COWAT, and Semantic Fluency however, produced consistent diagnostic rates, contributing further evidence for their validity.

Table 4. Percent of TBI and Community Sample Identified as Impaired Using Test Norms (Norm%) and Regression Equations (Reg%).

Test	TBI		Community		Regression Cut Off 10%ile
	Norm %	Reg %	Norm %	Reg %	
WCST	20%	17.5%	9%	9%	10.49
Perseverative Responses	22%	22%	7%	9%	8.75
Perseverative Errors	25%	22%	9%	9%	17.73
Total Errors	25%	22%	9%	9%	17.73
Semantic Fluency	18%	40%	4%	9%	7.10
TMT Part A Errors	32%	42%	6%	7%	12.01
TMT Part B Errors	57%	57%	6%	10%	20.84
COWAT Total	40%	40%	4%	12%	13.45
Design Fluency					
Free Novel Total	12%	0%	2%	4%	11.05
Fixed Novel Total	40%	15%	15%	8%	9.90

Note: BDI = Beck Depression Inventory; NART = National Adult Reading Test; WCST = Wisconsin Card Sorting Test; TMT = Trail Making Test; COWAT = Controlled Oral Word Association Test.

Discussion

The aim of the present study was to compute a series of regression equations in a healthy normal sample, and apply them to a group of patients with a history of moderate to severe TBI. Regression equations generate a predicted score that can be compared with the obtained score as an alternative to consulting normative tables, as an aid to the diagnosis of cognitive impairment. In effect, they allow the computation of an individualised normative score (the predicted score), and an estimate of error, based on demographic factors or other test scores (or some combination of both), which can be used to evaluate the abnormality of an obtained score. If the discrepancy is statistically abnormal, it may be an indication that the person tested has acquired brain damage. Because many tests have poorly developed norms or were standardised on samples quite different from the population the test is being used with (e.g., in terms of age, ethnicity, education, or intellectual ability), developing regression equations can be an efficient way of creating local norms.

A correlational analysis revealed that age, years of education, and NART error scores were modestly associated with at least some of the executive test scores. Most correlations were consistent with previous research (Crawford et al., 1992; Duncan, Johnson, Swales, &

Freer, 1997; Spreen & Strauss, 1998; Tombaugh et al., 1999). Multiple correlations between significant predictor variables and the executive test scores were about .50 for the WCST, TMT Part B, COWAT, and BCT. The amount of variance explained was comparable to those for similar equations used in the prediction of Wechsler IQ scores. For these tests (with the exception of the BCT, which was not considered) the percentage of individuals with regression-predicted test score discrepancies at or below the 10th percentile was comparable to the percentage scoring below the 10th percentile on the normative data. These results suggest that the regression equations are comparable in diagnostic sensitivity to the published normative data available for these measures, and provide good support for the use of the regression equations for these tests in clinical practice, as described below. Prediction of the TMT Part A, Semantic Fluency, and Design Fluency scores by the regression equations was less successful, with multiple correlations of the order of .30. The percentages of persons identified as impaired in the TBI group by the normative tables and the regression equations were more discrepant. In particular, the results cast doubt on the validity of the regression equations for the Design Fluency test. Use of published normative tables for this test is advisable.

It would be desirable to cross-validate the present findings with data from other studies, however, although a number of regression equations have been developed to predict IQ scores, few studies have published comparable equations to predict executive test scores. An exception is a study by Crawford et al. (1992) that used both demographic and NART data to predict premorbid scores on a test of verbal fluency. Using multiple regression analysis, they developed the following equation to predict scores on the COWAT, Total FAS score = 57.5 – (0.76 x NART), SEest = 9.09, which is very similar to the equation derived for the COWAT in the present study (Table 3). Crawford and colleagues identified a discrepancy between predicted and obtained of more than 12 as being impaired at the 10th percentile, whereas the equation from the present study identified a discrepancy of 13.5. A comparison of the slopes of theses two regression equations revealed no significant difference, $t(238) = .34, p = .75$ and chi-square analysis showed no difference in the number of persons identified as impaired using the two different equations.

As might be anticipated from previous research (Godfrey et al., 1993) there was a significant difference between the TBI and community groups on the BDI-II. Although many persons with TBI do become clinically depressed, this is not always the reason for elevated scores on a depression scale. In some cases, persons who have cognitive deficits consequent on a severe TBI may experience realistic feelings of loss and pessimism about the future, without necessarily being depressed. Further, some of the items on the BDI-II measure symptoms that could be a result of neurological damage rather than depression (e.g., concentration difficulty, fatigue). It is possible that level of depression, however, has a significant impact on test scores, indeed there were significant correlations between depression scores and TMT (Part A) and Semantic Fluency for the TBI group in the present investigation. Although it might be desirable to enter depression test scores as a predictor into a regression equation, this is difficult to achieve, because the correlations between cognitive performance and depression are typically low in a normal sample. Furthermore, depression is not a stable premorbid characteristic and as such, has no place in equations based on premorbid status. It is therefore necessary for clinicians to examine the individual item scores on a depression scale to determine whether scores on neuropsychological tests should be considered in the light of co-morbid depression.

Applying regression equations in clinical practice.

Although the use of regression equations did not appreciably affect the numbers of persons with TBI classified as impaired in the current study, the advantages of using regression equations to more precisely estimate an expected score and quantify the discrepancy between predicted and obtained scores are worth emphasising. A more precise specification of the abnormality of a discrepancy, based on regression equations developed on a local sample, may be of value in interpreting scores of from individual clients. To illustrate this, an example of the practical use of regression equations is given below. Before proceeding to this, however, consideration of some statistical and computer procedures for doing so presented by Crawford and colleagues (Crawford & Garthwaite, 2002; Crawford & Howell, 1998) is of relevance. Use of these methods adds appreciably to the accuracy with which a score discrepancy may be evaluated.

Traditionally, the method used to make inferences about the differences between obtained and predicted scores has been to multiply the SEest by an appropriate z-score in order to determine confidence intervals around the predicted score, and then to compute cut off scores. This is the way the results from the regression equations have been presented in Table 3. For example (referring to Table 3), any discrepancy between an individual's predicted WCST perseverative error score and the obtained score that exceeds 8.75, lies beyond the 90% confidence interval for the community sample and may be suggestive of impairment. Although this use of the SEest to assess score discrepancies is a useful approximation, it is technically incorrect. Crawford and colleagues (Crawford & Howell, 1998) observe that the more correct method of evaluating an individual's predicted score involves calculating the standard error for a new individual and multiplying this by the value of the t-statistic at the desired confidence limit. In practice, the traditional method and the method they propose, when compared on relevant data sets, produce comparable results, except where the sample size used to generate the equation is modest and the individual's score on the predictor variables is extreme. So using the data in Table 3 as a basis for interpreting score discrepancies is not likely to be misleading. Nonetheless, it is preferable to use the technically correct method and to facilitate this, Crawford has made available computer programs on the web, which can be downloaded and used with simple linear regression and multiple regression equations[1]. This makes it possible to use the data presented in this paper to test more precisely the abnormality of predicted and obtained score discrepancies.

The following example is provided to illustrate the use of Crawford's procedure for applying multiple regression weights to estimate test scores. Consider the situation where a clinician wishes to assess the abnormality of a score of 25 perseverative errors on the WCST made by a 40-year-old man with 12 years of education who makes 15 errors on the NART, and who has suffered a severe TBI six months previously. If the WCST manual is consulted it can be seen that such a score lies at the 9th percentile. Suppose, however, the clinician decided to apply the regression equation determined in the present study, on the grounds that it is based on New Zealand data and because it would give a more precise evaluation of the significance of the score than the norms, since the NART proved to be a better predictor of perseverative errors in combination with age than years of education (used in the published norms). The first step is to down load the relevant computer program as detailed in Footnote 1. Following the computer prompts, it is then necessary to enter the age (M = 39.10, SD = 11.75) and NART (M = 13.92, SD = 7.64) data for the sample (n = 100) for the two predictors of WSCT perseverative errors, together with the regression weights (Table 3; .42 for the NART, .28 for Age), and the correlation between NART and age (Table 2, r = -.22). Continuing to follow the prompts, the SEest for the regression equation (6.73) and the intercept is (-6.54) given in Table 3 are entered. Then the data for the client are entered: His age, NART score and obtained number of perseverative errors. When all these data have been entered, the results emerge: For a person of age 40 with a NART score of 15, at the 90% (1-tailed) confidence level, the program provides a SE value of 6.77 and a predicted score of 10.96. (Note that the new SE is similar to the value for the regression equation above, 6.73, which will be the case where the individual's predictor scores are close to those of the sample means). The 90% cut-off score would be 8.73, and the discrepancy between the predicted score and the obtained score is −14.04. The one-tailed probability of this man having a score this discrepant is equal to .02. A score of 25 errors for this 40-year-old man is statistically abnormal and may constitute evidence that a TBI has impaired his executive functioning.

In this case, the finding that the individual's score on the WCST was significantly discrepant was confirmed, and the degree of abnormality quantified in a way that is possibly more precise than using the published norms. The developers of the WCST did not have available NART scores for their sample, and correlational analyses in the present study suggested that NART errors were a better predictor of the WCST scores when combined with age than years of education. It is also likely that the percentage of people completing 12 years of education in the US is not the same as in New Zealand, making this a less reliable means of stratifying norms for clients outside the US. Thus where normative data from a local sample are available, use of regression equations to determine the abnormality of a discrepancy, at a minimum, can provide a useful means of validating conclusions based on the application of published norms collected in another country.

Conclusions

The construction of regression equations as an alternative to detailed normative tables is widespread in clinical neuropsychology. In the present study equations to predict scores on several measures of executive functioning were computed and applied to a sample of persons with TBI. The results were generally encouraging. Applying the equations gives a precise estimate of the abnormality of a discrepancy between obtained and expected scores in percentile terms, which is useful where tests have only limited cut-off points (e.g., BCT or TMT). For tests with more extensive published norms (e.g., WCST) the application of the equations, based on predictors relevant to the New Zealand setting may provide a better estimate of an individual's impairment. On occasion, it may also be useful to be able to validate results based on published US norms with results from a New Zealand sample. In general, the use of regression methods to develop norms has advantages in New Zealand, where limitations in available resources may make the collection of large and representative samples difficult.

References

Artiola i Fortuny, L., & Heaton, R. K. (1996). Standard versus computerized administration of the Wisconsin Card Sorting Test. *Clinical Neuropsychologist, 10*, 419-424.

Baddeley, A.D., & Wilson, B. (1988). Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition, 7*, 212-230.

Barona, A., Reynolds, C.R., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS-R. *Journal of Consulting and Clinical Psychology, 52*, 885-887.

Beck, A.T. (1996). *The Beck Depression Inventory II.* San Antonio: The Psychological Corporation, Harcourt, Brace and Company.

Borkowski, J.G., Benton, A.L., & Spreen, O. (1967). Word fluency and brain damage. *Neuropsychologia, 5*, 135-140.

Caradoc-Davies, T.H., & Dixon, G.S. (1995). Hospital morbidity due to head injuries in New Zealand in 1980-1988. *Neuroepidemiology, 14*, 199-208.

Cockburn, J. (1995). Performance on the Tower of London test after severe head injury. *Journal of the International Neuropsychological Society, 1*, 537-544.

Crawford, J.R. (1992). Current and premorbid intelligence measures in neuropsychological assessment. In Crawford, J.R., Parker, D.M., and McKinlay, W.W. (Eds.). *A handbook of neuropsychological assessment.* Hove: Erlbaum.

Crawford, J. R., & Allan, K. M. (1997). Estimating premorbid IQ with demographic variables: Regression equations derived from a U.K. sample. *The Clinical Neuropsychologist, 11*, 192-197.

Crawford, J.R., Allan, K.M., Cochrane, R.H., & Parker, D.M. (1990). Assessing the validity of NART-estimated premorbid IQs in the individual case. *British Journal of Clinical Psychology, 29*, 435-436.

Crawford, J.R., & Garthwaite, P.H. (2002). Investigation of the single case in neuropsychology: Confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia, 40,* 1196-1208.

Crawford, J.R., & Howell, D.C. (1998). Regression equations in clinical neuropsychology: An evaluation of statistical methods for comparing predicted and obtained scores. *Journal of Clinical and Experimental Neuropsychology, 20*, 755-762.

Crawford, J.R., Moore, J.W., & Cameron, I.M. (1992). Verbal fluency: A NART-based equation for estimation of premorbid performance. *British Journal of Clinical Psychology, 31*, 327-329.

Crawford, J. R., Obonsawin, M. C., & Allan, K. M. (1998). PASAT and components of WAIS-R performance: Convergent and discriminant validity. *Neuropsychological Rehabilitation, 8*, 255-272.

Crawford, J.R., Parker, D.M., & Besson, J.A.O. (1988). Estimation of premorbid intelligence in organic conditions. *British Journal of Psychiatry, 153*, 178-181.

Crawford, J.R., Parker, D.M., Stewart, L.E., Besson, J.A.O., & Delacey, G. (1989). Prediction of WAIS IQ with the National Adult Reading Test: Cross-validation and extension. *British Journal of Clinical Psychology, 28*, 267-273.

Crawford, J. R., Stewart, L. E., Parker, D. M., Besson, J. A. O., & Cochrane, R. (1989). Estimation of premorbid intelligence: Combining psychometric and demographic approaches improves predictive accuracy. *Personality and Individual Differences, 10*, 793-796.

Duncan, J., Johnson, R., Swales, M., & Freer, C. (1997). Frontal lobe deficits after head injury: Unity and diversity of function. *Cognitive Neuropsychology 14*, 713-741.

Franzen, M.D., Burgess, E.J., & Smith-Seemiller, L. (1989). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology, 12*, 711-738.

Freeman, J., & Godfrey, H.P.D. (2000). The validity of the NART-RSPM index in detecting intellectual decline following traumatic brain injury: A controlled study. *British Journal of Clinical Psychology, 39*, 95-103.

Glosser, G., & Goodglass, H. (1990). Disorders in executive control functions among aphasic and other brain damaged patients. *Journal of Clinical and Experimental Neuropsychology 12*, 485-501.

Godfrey, H.P.D., Partridge, F.M., Knight, R.G., & Bishara, S.N. (1993). Course of insight disorder and emotional dysfunction following severe closed head injury: A cross-sectional follow-up study. *Journal of Clinical and Experimental Neuropsychology, 15*, 503-515

Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd Ed.). Philadelphia: Lea and Febiger.

Heaton, R.K. (1981). *Wisconsin Card Sorting Test Manual.* Odessa, FL: Psychological Assessment Resources.

Heaton, R.K., Chelune, G.J., Tailey, J.L., Kay, G.G., & Curtis, G. (1993) *Wisconsin Card Sorting Test (WCST). Manual Revised and Expanded.* Odessa, FL: Psychological Assessment Resources.

Heaton, R.K. & PAR staff (1999). *Wisconsin Card Sorting Test: Computer Version 3 for Windows. Research Edition.* Odessa, Florida: Psychological Assessment Resources, Inc.

Heinrichs, R.W. (1990). Variables associated with Wisconsin Card Sorting Test performance in neuropsychiatric patients referred for assessment. *Neuropsychiatry, Neuropsychology and Behavioural Neurology, 3*, 107-112.

Hellman, S.G., Green, M.F., Kern, R.S., & Christenson, C.D. (1992). Comparison of card sorting and computer versions of the Wisconsin Card Sorting Test for psychiatric patients. *International Journal of Methods in Psychiatric Research, 2*, 151-155.

Johnstone, B., Leach, L.R., Hickey, M.L., Frank, R.G., & Rupright, J. (1995). Some objective measurements of frontal lobe deficits following traumatic brain injury. *Applied Neuropsychology, 2*, 24-28.

Jones-Gotman, M., & Milner, B. (1977). Design fluency: The intervention of nonsense drawing after focal cortical lesions. *Neuropsychologia, 15*, 653-674.

Karzmark, P., Heaton, R.K., Grant, I., & Mathews, C.G. (1985). Use of demographic variables to predict Full Scale IQ: A replication and extension. *Journal of Clinical and Experimental Neuropsychology, 7*, 412-420.

Krull, K., Scott, J., & Sherer, M. (1995). Estimation of premorbid intelligence from combined performance and demographic variables. *The Clinical Neuropsychologist, 9*, 83-87.

Martin, D.J., Oren, Z., and Boone, K. (1991). Major depressives' and dysthymics' performance on the Wisconsin Card Sorting Test. *Journal of Clinical Psychology, 47*, 685-690.

Mazaux, J.M., Masson, F., Levin, H.S., Alaoui, P., Maurette, P., & Barat, M. (1997). Long-term neuropsychological outcome and loss of social autonomy after traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 78*, 1316-1320.

Mercer, W.N., Harrell, E.H., Miller, D.C., Childs, H.W., & Rockers, D.M. (1997). Performance of brain-injured versus healthy adults in three versions of the Category Test. *Clinical Neuropsychologist, 11*, 174-179.

Milner, B. (1964). Some effects of frontal lobectomy in man. In Warren, J.M. & Akert, K. (Eds.) *The frontal granular cortex and behaviour.* New York: McGraw Hill. (pp. 313-334)

Moss, A.R., and Dowd, T. (1991). Does the NART hold after head injury? A case report. *British Journal of Clinical Psychology, 30*, 179-180.

Nelson, H.E., with Willison, J. (1991). *National Adult Reading Test (NART) Test Manual.* (2 edition.). Berkshire, England: NFER-NELSON Publishing Company Ltd.

O'Brien, L.F., Godfrey, H.P.D., Freeman, J., & Perkins, R. (1999). Determining clinically meaningful cognitive impairment following traumatic brain injury. *New Zealand Medical Journal, 112*, 295-297.

O'Carroll, R., Baikie, E.M., & Whittick, J.E. (1987). Does the National Adult Reading Test hold in dementia? *British Journal of Clinical Psychology, 26*, 315-316.

Raskin, S.A., & Rearick, E. (1996). Verbal fluency in individuals with mild traumatic brain injury. *Neuropsychology, 10*, 416-422.

Reitan, R.M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology, 19*, 393-394.

Russell, E.W., & Levy, M. (1987). Revision of the Halstead Category Test. *Journal of Consulting and Clinical Psychology, 55*, 898-901.

Shallice, T., & Burgess, P.W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114*, 727-741.

Spreen, O., & Benton, A.L. (1969). *Neurosensory Center Comprehensive Examination for Aphasia: Manual of Directions.* Victoria B.C: Neuropsychology Laboratory, University of Victoria.

Spreen, O., & Strauss, E. (1998). A *Compendium of Neuropsychological Tests: Administration, Norms and Commentary.* New York: Oxford University Press.

Stuss, D.T. & Benson, D.F. (1986). The frontal lobes. New York: Raven Press.

Tate, R.L. (1999). Executive dysfunction and characterological changes after traumatic brain injury: two side of the same coin? *Cortex, 35*, 39-55.

Tombaugh, T.N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology, 14*, 167-177.

Tranel, D., Anderson, S.W., & Benton, A. (1994). Development of "executive function" and its relationship to the frontal lobes. In Boller, F., and Grafman, J. (Eds.). *Handbook of Neuropsychology.* New York: Elsevier. pp 125-148.

Van den Broek, M.D., & Bradshaw, C.M. (1994). Detection of acquired deficits in general intelligence using the National Adult Reading Test and Raven's Standard Progressive Matrices. *British Journal of Clinical Psychology, 33*, 509-515.

Note

1. A range of programmes useful in the psychometric and statistical analysis, of particular relevance to single-case methodology in clinical psychology and neuropsychology, can be downloaded from: www.abdn.ac.uk/%7Epsy086/dept/psychom.htm. This page can also be located through the home page of Professor John Crawford of the Department of Psychology, University of Aberdeen. The programme used for applying regression data in the example in this paper is CLREGMUL.EXE. Data from a single predictor can be analysed with CLREGBIV.EXE. Note that when using these programmes, it is only necessary to enter the regression statistics once when analysing the results from more than one person on any occasion.

Acknowledgements

This research was completed with the assistance of Health Research Council Grant. The authors are grateful to the members of the Head Injury Societies and brain Injury Associations of Canterbury, Otago, Southland, Westland and Nelson, and community groups in Dunedin, for their participation in and assistance with the study. The comments of John Crawford on an earlier version of this paper are gratefully acknowledged. Mairead Harnett is now at the Department of Social and Preventive Medicine, University of Otago; Hamish Godfrey is deceased.

Address correspondence to:

Robert Knight, PhD
Department of Psychology
University of Otago
P.O. Box 56
Dunedin
Phone: +64 3 4797644
Fax: +64 3 4798335

Email: rknight@psy.otago.ac.nz