Journal of Clinical and Experimental Neuropsychology
1998, Vol. 20, No. 5, pp. 755-762
1380-3395/98/2005-755$12.00

Regression Equations in Clinical Neuropsychology: An Evaluation of Statistical Methods for Comparing Predicted and Obtained Scores*

J. R. Crawford[1] and David C. Howell[2]
[1]University of Aberdeen, UK and [2]University of Vermont, Burlington

ABSTRACT

Regression equations are widely used in clinical neuropsychology, particularly as an alternative to conventional normative data. In neuropsychological applications the most common method of making inferences concerning the difference between an individual's test score and the score predicted by a regression equation is to multiply the standard error of estimate by an appropriate value of *z* to form confidence limits around the predicted score. The technically correct method is to calculate the standard error of a new individual Y and multiply it by the value of *t* corresponding to the desired limits (e.g., 90% or 95%). These two methods are compared in data sets generated to be broadly representative of data sets used in clinical neuropsychology.

The former method produces confidence limits which are narrower than the true confidence limits and fail to reflect the fact that limits become wider as scores on the predictor deviate from the mean. However, for many of the example data sets studied, the differences between the two methods were trivial, thereby providing reassurance for those who use the former (technically incorrect) method. Despite this, it would be preferable to use the correct method particularly with equations derived from samples with modest *N*s, and for individuals with extreme scores on the predictor variable(s). To facilitate use of the correct method a computer program is made available for clinical practice.

Regression equations can serve a number of useful functions in clinical neuropsychology. Perhaps their most common role is as an alternative to the use of conventional normative data. For example, if it is found that age, years of education, and gender are related to performance on a memory test then a regression equation can be built (in a healthy sample) which uses these demographic variables as predictors. Thus, 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 (Crawford, 1996; Heaton, Grant, Ryan, & Matthews, 1996; Lezak, 1995).

Even with a single predictor, such as age, the regression approach is to be preferred over conventional normative data. It provides what Zachary and Gorsuch (1985) have termed ''continuous norms'' rather than the discrete norms formed by creating arbitrary age bands; in the latter case, the relative standing of an individual can change dramatically as they move from one band to another.

A second common application of regression is in the assessment of change in neuropsychological functioning in the individual case. Here, a regression equation can be built (normally using healthy participants) to predict an individual's level of performance on a neuropsychological instrument at retest from their score at initial testing. This approach simulta-

* Address correspondence to: John R. Crawford, Department of Psychology, King's College, University of Aberdeen, Aberdeen AB24 3HN, UK. E-mail: crawford@abdn.ac.uk.
Accepted for publication: July 27, 1998.

neously factors in the effects of practice (typically scores will be higher on retest) and regression to the mean (extreme scores on initial testing will, on average, be less extreme at retest).

In both of the foregoing examples, regression equations are used to generate a *predicted* or *expected* level of performance for an individual patient against which her/his *obtained* level of performance can be compared. A large discrepancy from the predicted score raises the suspicion of acquired cognitive deficit. The remainder of this paper is concerned with the statistical methods used to make inferences about the difference between an individual's predicted and obtained score. Before outlining the technically correct procedure, two methods commonly used in clinical neuropsychology will be described.

Using the Standard Error of Estimate ($S_{y.x}$) to Obtain a Confidence Interval

The standard error of estimate is a measure of the variability of observations about the regression line. As such, it reflects the precision of our estimation procedure. We can define the standard error of estimate ($S_{y.x}$) as

$$S_{y.x} = SD_y\sqrt{(1-r^2)\,\frac{N-1}{N-2}}, \qquad (1)$$

where SD_y = the standard deviation of the criterion variable, r^2 = the squared correlation between the predictor and criterion variables and $S_{y.x}$ = the sample size. A more familiar (approximate) formula for $S_{y.x}$ is given as

$$S_{y.x} = SD_y \quad \sqrt{1-r^2} \quad , \qquad (2)$$

and is a simplification of [1]. This formula yields a close approximation when the sample size is large, because (N-1)/(N-2) rapidly approaches one as sample size increases.

Both formulae for the standard error of estimate make it clear that the precision of our estimate when we take other variables into account ($S_{y.x}$) is greater than the precision when we look only at raw scores on the criterion variable, which is given by SD_Y. This difference increases as the correlation between the measure of interest and the predictor variable(s) increases. Without considering predictor variables, for large samples we would expect 90% of the sample to lie within 1.64 standard deviations of the mean of *Y*. Similarly, if we use *X* to predict *Y*, and if the assumptions underlying the use of regression are met, then one would expect that 90% of the sample in which a regression equation was built would lie within 1.64 $S_{y.x}$ units of the regression line.

It is important to note that this statement is cast in terms of the sample used to build the equation**.** However, researchers often present confidence limits based on $S_{y.x}$ as a means by which the clinician can evaluate the scores of individuals who were not in the original sample (Crawford, Moore, & Cameron, 1992; Knight & Shelton, 1983; McSweeny, Naugle, Chelune, & Lüders, 1993; Paolo, Ryan, Tröster, & Hilmer, 1996).

For example, McSweeny et al. (1993) built a regression equation in a sample of 50 patients with epilepsy to predict scores on the Wechsler Memory Scale – Revised (WMS-R, Wechsler, 1987) and Wechsler Adult Intelligence Scale – Revised (WAIS-R, Wechsler, 1981) from scores at initial testing. They created a 90% ''confidence interval'' by multiplying $S_{y.x}$ by 1.64 and recommended that the difference between an individual's predicted and obtained score should be considered 'significant' if it exceeded this interval (some authors would prefer the phrase ''prediction interval,'' or ''tolerance interval'' because the term ''confidence interval'' is generally reserved for intervals on parameters. However, in common with McSweeny et al., the present authors have chosen to use the term ''confidence interval'' because it is more commonly referenced in that way). Similarly, Knight and Shelton (1983) used data from previously published reports of repeat testing with the Wechsler Adult Intelligence Scale (WAIS, Wechsler, 1955) to generate regression equations. They multiplied $S_{y.x}$ by 1.64 and 1.96 to obtain what were referred to as the critical values required to achieve 'statistical significance' at the .05 and .01 levels (one-tailed) when comparing an individual's pre-

dicted retest score with her/his obtained retest score.

Tabulating Frequency Distribution of Discrepancies Between Predicted and Observed Scores

Another way of examining extreme deviations from predicted performance is to examine the frequency distribution of discrepancies between observed and predicted scores, tabulated for the sample used to build the equation. By referring to this table the clinician can observe the size of discrepancy required to exceed a given percentage of the sample. Provided that the assumptions underlying the use of regression are met, and the original sample is sufficiently large, this approach will yield 'critical values' that closely resemble those produced by the foregoing method; in the former case, the rarity of a discrepancy is estimated statistically, in the present case, it is empirically derived.

This technique has been commonly used with methods designed to estimate premorbid intelligence; for reviews see Crawford and O'Carroll. For example, Nelson (1982) built a regression equation using the National Adult Reading Test (NART, Nelson, 1982) to estimate premorbid WAIS IQs. The NART manual presents a table which records the frequency distribution in the standardization sample of discrepancies between obtained WAIS scores and scores predicted from the NART. The same approach was used when the NART was standardised against the WAIS-R (Nelson & Willison, 1991) and has also been used with approaches to the estimation of premorbid ability that are based on demographic variables (see Crawford & Allan, 1997).

Using Standard Error of a New Individual $\hat{Y}$ to Form Confidence Intervals

The foregoing methods take account of error in predicting an individual's score from the regression of the dependent variable on the independent variable(s), and are, therefore, a reasonable way of drawing inferences concerning the score of an individual drawn from the sample used to generate the equation. However, they ignore the error arising from using sample regression coefficients to estimate population regression coefficients. This has two consequences; first, the confidence limits will be too narrow and second, the true confidence limits will be wider as the score on the predictor deviates from the predictor mean. For an intuitive understanding of this latter effect, consider what would happen to the predictions for different values of X if we rotated the regression line slightly around the mean of X. A further issue is that, in the earlier examples of the use of the $S_{y.x}$ method, $S_{y.x}$ was multiplied by a value of z rather than t (with df of $N - 2$) The t statistic should be used because the estimates are based on sample statistics rather than parameters. Although this will have a negligible effect with large samples, it constitutes another factor which will produce an underestimation of the true confidence limits.

The correct method of drawing inferences concerning an individuals' predicted score is to use the standard error of a new individual $\hat{Y}$ ($SE_{yo-\hat{Y}o}$) to form confidence limits on Y (Howell, 1997). When there is a single predictor, a formula for the $SE_{yo-\hat{Y}o}$ is as follows:

$$SE_{yo-\hat{Y}o} = S_{y.x}\sqrt{1 + \frac{1}{N} + \frac{(X_o - \bar{X})^2}{SD_x^2(N-1)}}, \quad (3)$$

where $S_{y.x}$ is as defined previously, $\bar{X}$ = the mean score on the predictor variable, X_o = the individual's score on the predictor, and SD_x = the standard deviation of the predictor variable (there are a number of equivalent formula for $SE_{yo-\hat{Y}o}$, the present one is used because it is the simplest). Having calculated the $SE_{yo-\hat{Y}o}$, confidence limits on Y can be obtained from the following formula:

$$CI(Y) = \hat{Y} \pm (t_{\alpha/2})\,(SE_{yo-\hat{Y}o}, \quad (4)$$

where $t_{\alpha/2}$= the value of t (with $df = N - 2$) corresponding to the required confidence limits. Thus, as a simple example, if $N = 52$, $SE_{yo-\hat{Y}o}$ = 10, and 90% (two-sided) confidence limits were required, 10 would be multiplied by 1.676. If an individual's predicted score ($\hat{Y}$) was 50 then the 90% upper confidence limit for Y would be 66.76 and the lower limit would be 33.24.

To our knowledge no study in clinical neuropsychology has adopted this latter approach. In contrast, statistical, psychometric, and biometric texts advocate it consistently (Armitage & Berry, 1994; Cohen & Cohen, 1983; Daly, Hand, Jones, Lunn, & McConway, 1995; Daniel, 1983; Draper & Smith, 1981; Gardner & Altman, 1989; Howell, 1997; Pedhazur, 1982; Sokal & Rohlf, 1981; Wittink, 1988; Zar, 1984, 1996). Moreover, some of these texts explicitly warn against using confidence intervals based on $S_{y.x}$. Zar (1984), for example, urges researchers to avoid the practice of following a regression equation with ''$\pm S_{y.x}$'', in order to, ''keep the reader from inferring that this is an equation for determining a confidence interval for $\hat{Y}$'' (p. 271).

The present authors do not dispute that the above procedure is the correct one and that the two commonly employed methods are technically incorrect. However, it is legitimate to explore the extent to which the choice of method makes a difference in *practical terms*. Many of the graphical representations used to depict the elliptical nature of confidence intervals suggest that the methods will yield very divergent results. For example, Figure 1 is adapted from Wittink (1988) and implies a dramatic difference in the magnitude of the confidence interval as scores on the predictor variable diverge from their mean. However, there may be an element of exaggeration for effect in such representations.

Comparison of Confidence Intervals Based on $S_{y.x}$ and $SE_{yo-\hat{y}o}$

In this section the methods are compared in hypothetical data sets designed to be reasonably typical of those found in clinical neuropsychology in terms of sample size, the precision of measurement of the criterion variable, and the explanatory power of the predictor variable(s). For simplicity, the illustrative data sets are all limited to the case of one predictor variable and scores on the criterion and predictor variables are expressed as *T* scores. *T* scores were selected because they are familiar to clinicians and are often used when test scores are converted to a common metric (Crawford, 1996; Lezak, 1995). Additionally, most scales used in clinical neuropsychology express scores as integers with only two or three significant digits, for example, percentiles, Wechsler subtest scaled scores, IQs, and Memory Quotients (MQs). Therefore, there is no loss in generality from using *T* scores for these calculations.

The sample sizes of the data sets vary from 25 to 500. An *N* of 25 may seem modest but samples of this size, and smaller, have been used to generate equations for clinical use (Knight & Shelton, 1983). It can be seen from formula (1) that $S_{y.x}$ shrinks as sample size increases. Therefore, sample size will exert an effect on both methods as both incorporate $S_{y.x}$ in their computations. However, sample size will exert an additional, unique, effect on $SE_{yo-\hat{y}o}$ because the error in estimating the population regression coefficient will shrink as sample size increases.

For each of the selected sample sizes, the results from both methods are compared for two values of the correlation (r_{xy}) between predictor and criterion (0.50 and 0.75). Because the predictor and criterion variables have the same mean and *SD* in these examples, the regression coefficient (*b*) takes the same value as the corresponding $r_{xy.}$ These values of r_{xy} are not atypical in neuropsychological applications. Although regression equations based on data with an *r* (or multiple *R*) below 0.5 would still have advantages over conventional normative data, such equations are rarely published or used. Correlations above .75 are less rare, particularly when scores on initial testing are used to predict scores at retest. However, preparatory work for this paper revealed that the results for higher *r*s (e.g., $r = .85$) were similar to those for $r = .75$. Therefore, in the interests of brevity such data are not presented.

Table 1 records 90% (two-sided) confidence intervals generated by applying the common and correct methods. The third column of Table 1 records $S_{y.x}$; multiplying $S_{y.x}$ by 1.64 produced the width of the confidence intervals in column four. The remaining columns record the widths of the confidence intervals generated from $SE_{yo-\hat{y}o}$ and critical values from the *t* distribution. As noted, these confidence intervals become wider as scores on the predictor vari-

able(s) diverge from the mean of the predictor variable. Confidence interval widths are presented for scores which are 0, 0.5, 1, 1.5, 2, 3, and 4 *SD* units from the mean; as scores on the predictor variables are expressed as *T* scores, these correspond to *T* scores of 50, 55, 60, 65, 70, 80, and 90. (We have chosen to present scores which are above the mean, however the confidence interval widths would be identical for scores which were the equivalent number of *SD* units below the mean). The predicted scores from the regression equations when $b = 0.5$ and $b = 0.75$ are presented above their corresponding confidence intervals; these predicted scores are rounded to the nearest *T* score. To illustrate this table, take the specific example of a regression equation based on a sample with an *N* of 50 and an *r* and *b* of 0.5. If an individual's score on the predictor variable was 65 then it can be seen that the predicted score ($\hat{Y}$) would be 58. The 90% confidence limits based on $SE_{yo-\hat{y}o}$ in this example is 58 +15.15 (15 when rounded to the nearest *T* score). Thus, if the discrepancy between an individual's obtained score on the criterion variable and the predicted score exceeded 15, it would exceed the 90% confidence limits, that is, the probability is less than 10% that this size of discrepancy would arise by chance. Alternatively, if the $S_{y.x}$ were used to construct the confidence limits, then the discrepancy would have to exceed 14.35 (14 when rounded to the nearest *T* score).

Table 1 demonstrates that the confidence limits based on $S_{y.x}$ underestimate those generated by the correct method for *all* values of the predictor variable. It can also be seen that the magnitude of the difference between these confidence intervals increases as the score on the predictor deviates from the mean. However, a notable feature of the present examples is that, for many of the entries in Table 1, the difference between these two methods is modest. For example, if the intervals were rounded to the nearest *T* score, then in many cases the two methods either yield the same interval or differ by only one *T*-score unit. Thus, in the present examples, the differences between the confidence intervals based on $S_{y.x}$ and $SE_{yo-\hat{y}o}$ only become substan-

Table 1. Comparison of 90% (Two-Sided) Confidence Intervals[a].

				90% Confidence Interval for $\hat{Y}$ based on $SE_{yo-\hat{y}o}$						
b / *r*	*N*	$S_{y.x}$	$S_{y.x}$ 90%CI	X = 50	X = 55	X = 60	X = 65	X = 70	X = 80	X = 90
b = 0.50				$\hat{Y}$ = 50	$\hat{Y}$ = 53	$\hat{Y}$ = 55	$\hat{Y}$ = 58	$\hat{Y}$ = 60	$\hat{Y}$ = 65	$\hat{Y}$ = 70
	25	8.85	14.51	15.46	15.54	15.77	16.15	16.66	18.04	19.81
	50	8.75	14.35	14.82	14.86	14.97	15.15	15.40	16.10	17.03
	100	8.70	14.28	14.53	14.54	14.60	14.69	14.81	15.17	15.65
	200	8.68	14.24	14.38	14.39	14.42	14.46	14.53	14.70	14.95
	500	8.67	14.22	14.30	14.30	14.31	14.33	14.36	14.43	14.53
b = 0.75				$\hat{Y}$ = 50	$\hat{Y}$ = 54	$\hat{Y}$ = 58	$\hat{Y}$ = 61	$\hat{Y}$ = 65	$\hat{Y}$ = 73	$\hat{Y}$ = 80
	25	6.76	11.08	11.81	11.87	12.04	12.33	12.72	13.78	15.13
	50	6.68	10.96	11.32	11.35	11.43	11.57	11.77	12.30	13.00
	100	6.65	10.90	11.09	11.11	11.15	11.22	11.31	11.58	11.95
	200	6.63	10.88	10.99	10.99	11.01	11.05	11.10	11.23	11.42
	500	6.62	10.86	10.92	10.93	10.95	10.97	10.99	11.02	11.10

Note. [a]The confidence intervals were obtained from $S_{y.x}$ and from $SE_{yo-\hat{y}o}$ for varying sample sizes and two values of *r* / *b* (the examples are based on predictor and criterion variables that are expressed as *T* scores). Confidence intervals based on $S_{y.x}$ do not vary as a function of scores on the predictor and are presented in column four. Confidence intervals based on the use of $SE_{yo-\hat{y}o}$ are presented in the subsequent columns for each of seven values of the predictor (X), ranging from a predictor score that is at the mean (X = 50), to a score that is very extreme (X = 90). The *predicted* scores ($\hat{Y}$) are presented above each of the two sets of confidence intervals; the two sets of confidence intervals and predicted scores differ as a function of the correlation (*r*) / regression coefficient (*b*).

tial when the sample size is relatively modest ($\leq$ 100) and an individual's score on the predictor variable is extreme ($\geq$ 2 *SD*s). In practice, it would often be inappropriate to use a regression equation when these circumstances hold, because the individual's score on *X* may lie beyond the range of the values of *X* in the sample used to create the regression equation (the relationship between predictor and criterion may become nonlinear for scores that lie beyond the range of scores observed in the sample). We occasionally have to make such predictions, but considerable care should be taken in their interpretation, and particularly in the interpretation of a substantial deviation from this prediction.

The illustrative data were selected to be reasonably typical of those found in clinical neuropsychology in terms of sample size, the precision of measurement of the criterion variable, and the explanatory power of the predictor variable. Thus, the results provide some reassurance for those clinicians who use confidence limits based on $S_{y.x}$ when drawing inferences regarding an individual's score. As noted, the use of tabulated frequency distributions of discrepancies between observed and predicted scores will normally yield very similar results to the use of confidence limits based on $S_{y.x}$; therefore, the present data also provide reassurance for those using this latter method.

The present observations appear to be at odds with the conclusion that could be drawn from inspection of many of the graphical representations of the elliptical nature of confidence intervals (i.e., see Fig. 1). The reason for this apparently dramatic disparity arises because many graphical representations do not provide a scale for the abscissa (*X* axis). It may not be inappropriate to exclude a scale, given that such illustrations are intended to be generic. However, the result is that they obscure the fact that confidence intervals based on $SE_{yo\text{-}\hat{y}o}$ become *heavily* elliptical only at very extreme values of the predictor variable (i.e., > 4 *SD*s). Such extreme values are rare with neuropsychological data and, as noted, when they occur often preclude (or make problematic) the use of a regression prediction for that individual in any case.

Computer Programs for Confidence Intervals Based on $SE_{yo\text{-}\hat{y}o}$

The present examples suggest that, in many circumstances, substituting confidence intervals based on $S_{y.x}$ for the correct method does not

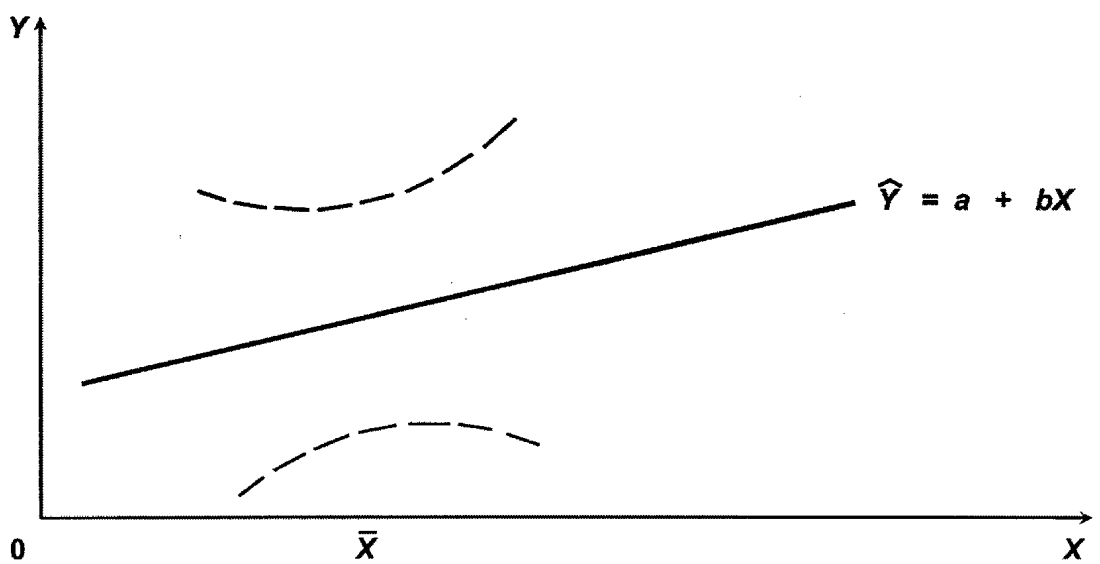

Fig. 1. A generic illustration of the elliptical nature of confidence intervals for an individual's score based on an illustration by Wittink (1988).

seriously compromise the validity of inferences concerning an individual's score. However, it remains the case that the use of $S_{y.x}$ to generate confidence limits is technically incorrect and will *systematically* yield confidence limits which are *narrower* than those obtained by using the correct method. Therefore, it would be preferable to use the correct method, particularly when sample size is relatively modest and scores are relatively extreme. On the other hand, it is undeniably more complicated and time-consuming to generate confidence limits using $SE_{yo\text{-}\hat{y}o}$ than $S_{y.x}$. To address this, the first author has written a computer program for PCs that automates the process. The program prompts the user for *b* (the regression coefficient), *a* (the intercept), the mean and *SD* of the predictor variables, and $S_{y.x}$. This information can be saved to file for future use. The user then enters an individual's score on the predictor and the desired confidence interval (e.g., 85% or 95% etc.). The output consists of the predicted score and confidence limits.

In clinical practice there will often be a need for *one*-sided confidence intervals. For example, if a regression equation is being used in place of conventional norms, the clinician may have hypothesised that an individual's obtained score will be significantly below that predicted by the equation (i.e., the hypothesis is directional). To deal with this situation, the program can generate one-sided intervals; it will be appreciated that this option should not be used post hoc.

Aside from saving time, use of this program reduces the chance of clerical error. Furthermore, it can also provide an exact probability (one- or two-tailed) for the discrepancy between an individual's predicted and obtained scores. This is achieved by dividing the discrepancy between an individual's predicted and obtained score by $SE_{yo\text{-}\hat{y}o}$. This yields a *t* statistic with degrees of freedom equal to $N - 2$ for which an exact *p* is then calculated (i.e., in essence this procedure simply rearranges the terms in the formula for a confidence interval (4) to solve for *t*).

In the interests of simplicity the examples and related discussion have, to this point, been limited to the case of a single predictor variable. However, the authors have also written a computer program to generate confidence limits for use with multiple regression equations. The computations are considerably more complicated when there are multiple predictor variables, to the extent that it is not realistic to generate them by hand or calculator. This program prompts for *a* (the intercept), the regression coefficients (*b*s), the means and *SD*s of the predictor variables, and $S_{y.x}$. The intercorrelations of the predictor variables is also required. After entering these data they can be saved to file for future use. Finally, an individual's scores on the predictor variables are entered and the required confidence interval specified. The program transforms the individual's scores on the predictors to *z* scores and inverts the correlation matrix. The $SE_{yo\text{-}\hat{y}o}$ is then obtained by implementing the following formula, which, with minor changes in notation, is taken from Cohen and Cohen (1983):

$$SE_{yo\text{-}\hat{y}o} = \frac{S_{y.x}}{\sqrt{n}}\sqrt{n + 1 + \Sigma r^{ii} z_{io}^2 + 2\Sigma r^{ij} z_{io} z_{jo}} \qquad (5)$$

where r_{ij} identifies off-diagonal elements of the inverted correlation matrix, r_{ij}identifies elements in the main diagonal, and *z* identifies the individual's scores on the predictor variables in *z*-score form. The first summation is over the *k* diagonal elements and the second is over the *k(k–1)/2* off-diagonal elements below (or above) the diagonal. This program offers the same options as those outlined for the case of a single predictor.

Although the APA Publication Manual (American Psychological Association, 1994) recommends that authors reporting multiple regression analyses should include the correlation matrix as a table, this is not always done. Therefore, it will not always be possible to use this latter program to generate confidence intervals. Compiled versions of the two programs can be downloaded from the first author's website at the following address: http://www.psyc.abdn.ac.uk/homedir/jcrawford/clreg.htm.

REFERENCES

American Psychological Association (1994). *Publication manual of the American Psychological Association* (4th ed.) Washington, DC: Author.

Armitage, P., & Berry, G. (1994). *Statistical methods in medical research*. Oxford, UK: Blackwell.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Crawford, J. R. (1992). Current and premorbid intelligence measures in neuropsychological assessment. In J. R. Crawford, D. M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (pp. 21-49). London: Lawrence Erlbaum.

Crawford, J. R. (1996). Assessment. In J. G. Beaumont, P. M. Kenealy, & M. J. Rogers (Eds.), *The Blackwell dictionary of neuropsychology* (pp. 108-116). London: Blackwell.

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

Daly, F., Hand, D. J., Jones, M., C., Lunn, A. D., & McConway, K. J. (1995). *Elements of statistics*. Wokingham, UK: Addison-Wesley.

Daniel, W. W. (1983). *Biostatistics: A foundation for analysis in the health sciences*. New York: Wiley.

Draper, N. R., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). New York: Wiley.

Gardner, M. J., & Altman, D. G. (1989). *Statistics with confidence – confidence intervals and statistical guidelines*. London: British Medical Journal.

Heaton, R. K., Grant, I., Ryan, L., & Matthews, C. G. (1996). Demographic influences on neuropsychological test performance. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (2nd ed.) (pp. 141-163). New York: Oxford University Press.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury Press.

Knight, R. G., & Shelton, E. J. (1983). Tables for evaluating predicted retest changes in Wechsler Adult Intelligence Scale scores. *British Journal of Clinical Psychology, 22*, 77-81.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

McSweeny, A. J., Naugle, R. I., Chelune, G. J., & Lüders, H. (1993). ''T scores for change'': An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist, 7*, 300-312.

Nelson, H. E. (1982). *National Adult Reading Test (NART): Test manual*. Windsor, UK: NFER -Nelson.

Nelson, H. E., & Willison, J. (1991). *National Adult Reading Test manual* (2nd ed.). Windsor, UK: NFER-Nelson.

O'Carroll, R. (1995). The assessment of premorbid ability: A critical review. *Neurocase, 1*, 83-89.

Paolo, A. M., Ryan, J. J., Troster, A. I., & Hilmer, C. D. (1996). Demographically based regression equations to estimate WAIS-R subtest scaled scores. *The Clinical Neuropsychologist, 10*, 130-140.

Pedhazur, E. J. (1982). *Multiple regression in behavioural research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart & Winston.

Sokal, R. R., & Rohlf, J. F. (1981). *Biometry*. San Francisco, CA: W.H. Freeman.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.

Wechsler, D. (1987). *Manual for the Wechsler Memory Scale-Revised*. San Antonio, TX: Psychological Corporation.

Wittink, D. R. (1988). *The application of regression anlaysis*. Boston: Allyn & Bacon.

Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology, 41*, 86-94.

Zar, J. H. (1984). *Biostatistical analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Zar, J. H. (1996). *Biostatistical analysis* (3rd ed.). London: Prentice-Hall.