Journal of the International Neuropsychological Society (2005), **11**, 788–789.

DOI: 10.1017/S1355617705050940

LETTER TO THE EDITOR

Methodology is more important than statistics when determining reliable change

ANTON D. HINTON-BAYRÉ
School of Medicine, University of Queensland, Herston, Queensland, Australia
(Received June 15, 2005; Accepted July 25, 2005)

McCrea and colleagues (2005) sought to evaluate the efficacy of a multidimensional assessment approach to the management of concussion in contact sport in their article entitled "Standard regression-based methods for measuring recovery after sport-related concussion." They used the Standardized Regression-Based (SRB) method to determine when an athlete was recovered. Though not in the above article, the SRB is often presented alongside the Reliable Change Index adjusted for practice (RCIp) and the clinician must choose which to use with limited direction. This letter seeks to further elaborate on why the SRB may be preferred and how tighter methodology is perhaps even more important. The general form for any reliable change can be written as:

$$\mathrm{RC} = \frac{(\text{Actual retest score } [Y] - \text{Predicted retest score } [Y'])}{\text{Standard error } [SE]}.$$

RCIp makes a constant adjustment for practice and assumes error is also uniform across individuals. Under RCIp, Y' is simply the pretest score plus a mean practice, and the error is the *SE* of measurement of the difference, based on retest comparison data (Chelune et al., 1993). In contrast, SRB effectively adjusts both the practice effect and the error term according to the individual's score relative to the mean of the comparison group on the predictors. Using SRB, Y' is adjusted by the mean practice effect and regression to the mean. Relative to RCIp, SRB adjusts low scorers upwards and high scorers downwards towards the mean on retesting. Classically, the error for an individual also depends on their relative standing on predictors, with more extreme scores having a greater margin of error (Pedhazur, 1997). Several authors suggest using the Standard Error of Prediction (SE_P) as the error term for SRB. The SE_P used often referred to the standard deviation of residuals, a constant value not classically used for regression prediction. In many studies it is not clear which error term was used. A good example of using individualized error is provided by Salinsky and colleagues (2001). The individualized SE_P formula can be found in several basic statistics texts (e.g., Pedhauzur, 1997, p. 205).

However, does the further advance in statistical sophistication from RCIp to multiple SRB improve the accuracy of clinical decision-making? Several authors have demonstrated that false positive rates for RCIp and multiple SRB do not differ substantively (e.g., Temkin et al., 1999). Moreover, these researchers have found that outside the pretest score, other predictors add little if any improvement in accuracy. RCIp and multiple SRB will provide comparable false positive rates when using reliable measurement, as regression to the mean will usually be negligible. Thus, the two methods usually also agree in terms of who has "reliably changed." This cannot be clearly said when measurement is less reliable. The exceptions will occur in those at the more extreme ends of the predictor range. Clinically, if one uses a measure with poor reliability and RCIp, they will be less likely to classify low scorers as impaired and more likely to classify high scorers as impaired, compared to SRB—this assumes lower scores = poorer performance. I agree the multiple SRB has a theoretical advantage. However, the present author has become less concerned about the type of reliable change method employed and more concerned that reliable measures be used in a reliable way ($r > .70$), and that comparison retest norms fit the clinical individual closely (± 1 *SD*).

Reprint requests to: Anton D. Hinton-Bayré, Ph.D., Cognitive Psychophysiology Laboratory, School of Medicine, University of Queensland, Herston Road, Herston, Queensland, 4064 Australia. E-mail: s309339@student.uq.edu.au

REFERENCES

Chelune, G.J., Naugle, R.I., Lüders, H., Sedlak, J., & Awad, I.A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology*, *7*, 41–52.

McCrea, M., Barr, W.B., Guskiewicz, K., Randolf, C., Marshall, S.W., Cantu, R.C., Onate, J.A., & Kelly, J.P. (2005). Standard regression-based methods for measuring recovery after sport-related concussion. *Journal of the International Neuropsychological Society*, *11*, 58–69.

Pedhazur, E.J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fortworth, TX: Harcourt Brace.

Salinsky, M.C., Storzbach, D., Dodrill, C.B., & Binder, L.M. (2001). Test–retest bias, reliability, and regression equations for neuropsychological measures repeated over a 12–16-week period. *Journal of the International Neuropsychological Society*, *7*, 597–605.

Temkin, N.R., Heaton, R.K., Grant, I., & Dikmen, S.S. (1999). Detecting change in neuropsychological test performance: A comparison of four models. *Journal of the International Neuropsychological Society*, *5*, 357–369.

https://doi.org/10.1017/S1355617705050940 Published online by Cambridge University Press