

# Response to Editorial note on: Agreement of repeated motor and sensory scores at individual myotomes and dermatomes in young persons with spinal cord injury

*Spinal Cord* (2013) **51**, 83; doi:10.1038/sc.2012.130; published online 6 November 2012

**Editorial Note on:** *Spinal Cord* (2013) **51**, 82; doi:10.1038/sc.2012.129; published online 30 October 2012

To address the first concern, the authors do agree that analyzing these data at each individual spinal level or segmental groups (C1-C4, C5 T1 and so on) would result in very useful information. However, we do not have enough data at each level or when grouped by levels to achieve reasonable statistical power.

Second, although it is true that the intraclass correlation coefficient (ICC) uses inter-subject variability, it is not a measure of it. It relates the variability among raters (that is, agreement) to the inter-subject variability in the form of a ratio. We believe this is a more appropriate way of 'adjusting' the absolute agreement than that used by the kappa statistic. The kappa statistic uses an arbitrary 'chance agreement' formula to adjust the absolute agreement regardless of the individual experimental variability.

Finally, it is well established that the kappa coefficient and the intra-class correlation are equivalent indices for inter- and intra-rater reliability. Although kappa is sometimes preferred for simple experimental designs, an analysis of variance (ANOVA) for repeated measures was necessary for this experiment. Repeated measures of each patient require the estimation of the within-subject covariance structure. Kappa could not accommodate this design. The analysis of repeated measures of nominal and ordinal data in this experiment further required the transformation of the data into normalized ranks before analysis. In this way, it was possible to use ANOVA methodology for categorical data. Please see references below.

Fleiss JL, Cohen J. The equivalence of weighted kappa and the intra-class correlation as measures of reliability. *Educ Psychol Measurement* 1973; **33**: 613–619.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420–428.

Koch GG. A general methodology for the analysis of experiments with repeated measures of categorical data. *Biometrics* 1977; **33**: 133–158.

Ridout MS, Demetrio CGB, Firth D. Estimating intraclass correlation for binary data. *Biometrics* 1999; **55**: 137–148.

Conover WJ, Iman RL. Rank transformations as a bridge between parametric and nonparametric statistics. *Am Statistic* 1981; **35**: 124–129.

Harter HL. Expected values of normal order statistics. *Biometrika* 1961; **48**: 151–165.

## CONFLICT OF INTEREST

The author declares no conflict of interest.

L Krisa<sup>1</sup>, J Gaughan<sup>2</sup>, L Vogel<sup>3,4</sup>, RR Betz<sup>1</sup> and MJ Mulcahey<sup>1,5</sup>

<sup>1</sup>Shriners Hospitals for Children, Philadelphia, PA, USA; <sup>2</sup>Temple University School of Medicine, Biostatistical Consulting Center, Philadelphia, PA, USA; <sup>3</sup>Shriners Hospitals for Children, Chicago, IL, USA; <sup>4</sup>Rush Medical College, Chicago, IL, USA and <sup>5</sup>Department of Occupational Therapy, Jefferson School of Health Professions, Thomas Jefferson University, Philadelphia, PA, USA  
E-mail: Maryjane.mulcahey-hershey@jefferson.edu