

## CORRESPONDENCE

# Response to: Reliability Of the International Spinal Cord Injury Musculoskeletal Basic Data Set; Methodological and Statistical Issue to Avoid Misinterpretation

*Spinal Cord Series and Cases* (2016) 2, 16024; doi:10.1038/scsandc.2016.24; published online 29 September 2016

We thank Sabour and Ghassemi<sup>1</sup> for their comments to our article,<sup>2</sup> and contribution to improve the quality of reliability studies through choosing the right statistical measures.

First issue raised<sup>1</sup> is the use of Kappa statistics as a measure of reliability (agreement, precision) and weighted Kappa. Weighted Kappa requires the assessed variables to be of ordinal nature. However, the outcome variables in our study were mainly dichotomous and thus a weighting scheme was not possible.<sup>3</sup> Only the last variable in the International Spinal Cord Injury Musculoskeletal Basic Data Set (ISCIMSBDS),<sup>4</sup> with the question: 'Do any of the above musculoskeletal challenges interfere with your activities of daily living (transfers, walking, dressing, showers, and so on.?)' and the options: 'not at all', 'yes a little', 'yes a lot' is ordinal. One could, perhaps, argue that variables concerning 'Fractures', 'Heterotopic ossifications', 'Contractures' and 'Degenerative Changes/Overuse', seen in the table of the ISCIMSBDS can be considered ordinal, concerning the locations of the above mentioned variables. In this case, two raters choosing adjacent locations would be considered as better agreement than locations physically longer apart. For example, in the case where the first rater choose the location 'Elbow' and the second rater choose 'Shoulder/Humerus', then this would reflect better agreement than in the case where the second rater chose the location 'Foot'. This could be a reasonable argument, but from a clinical perspective we found it more appropriate and relevant only to consider exact agreement of locations as measure of reliability.

Two weaknesses of Kappa were mentioned. Firstly the fact that Kappa is affected by prevalence, exemplified in Figure 1 by Sabour and Ghassemi.<sup>1</sup> A skewed distribution between the two concordant pairs results in a lower kappa value (Figure 1a) despite having the same percentage/crude agreement of the two concordant and discordant pairs. We agree with this concern and was also mentioned in the statistics section of the article.<sup>2</sup> The fact that Kappa is sensitive to prevalence can therefore be seen as a limitation. It could, however, also be argued to be an advantage. Kappa is chance-corrected agreement or agreement beyond chance. In a population with the prevalence seen in Figure 1a, there would be an increased probability of raters to agree by chance, and Kappa adjust for this. Sim and Wright<sup>5</sup> use the Prevalence Index to describe this effect:  $\frac{a-d}{n}$  (see Table 1). A high Prevalence Index is followed by a low Kappa and vice versa. We argue that this can be a desired property of Kappa, but that the prevalence should be taken into account when interpreting the Kappa value. This is the reason why we reported the prevalence of the symptoms in our article as well as reporting the percentage agreement.

The second mentioned limitation of Kappa is that it depends upon the number of categories, which means the more categories, the lower the kappa value. This is true when using the weighted Kappa, but does not apply to the present study as mentioned earlier.

Finally, Sabour and Ghassemi address the importance of having an individual-based approach instead of a group based. We can

**Table 1.** A 2×2 contingency table of agreement between two observers with concordant pairs (a, d) and discordant pairs (b, c)

	Observer 1	
	Positive	Negative
Observer 2 Positive	a	b
Negative	c	d

confirm that for both intra- and inter-rater evaluations, all comparisons were performed pair-wise between the two relevant ratings, either performed by the same rater twice (intra-rater) or by two different raters (inter-rater), and then summed in a 2×2 contingency table used for both percentage (crude) agreement and Kappa calculations.

## COMPETING INTERESTS

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank Karl Bang Christensen, Department of Public Health, Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark, for statistical counselling.

Carsten B Baunsgaard<sup>1</sup>, Harvinder S Chhabra<sup>2</sup>, Lisa A Harvey<sup>3</sup>, Gordana Savic<sup>4</sup>, Sue Ann Sisto<sup>5</sup>, Faiza Qureshi<sup>5</sup>, Gaurav Sachdev<sup>2</sup>, Mofid Saif<sup>4</sup>, Rajesh Sharawat<sup>2</sup>, Jayne Yeomans<sup>6</sup> and Fin Biering-Sørensen<sup>1</sup>

<sup>1</sup>*Clinic for Spinal Cord Injuries, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark;*

<sup>2</sup>*Indian Spinal Injuries Centre, New Delhi, India;*

<sup>3</sup>*John Walsh Centre for Rehabilitation Research, Kolling Institute, Sydney Medical School Northern, University of Sydney, Sydney, NSW, Australia;*

<sup>4</sup>*National Spinal Injuries Centre, Stoke Mandeville Hospital, Buckinghamshire Healthcare NHS Trust, Aylesbury, UK;*

<sup>5</sup>*School of Health Technology & Management, Research and Development Park, Rehabilitation Research and Movement Performance (RRAMP) Laboratory, Stony Brook University, New York, NY, USA and*

<sup>6</sup>*Spinal Injury Unit, Royal North Shore Hospital, Sydney, Australia*

Correspondence: CB Baunsgaard  
(carsten.bach.baunsgaard@regionh.dk or  
cbbunsgaard@sund.ku.dk)

## REFERENCES

- 1 Sabour S, Ghassemi F. Reliability of the International Spinal Cord Injury Musculoskeletal Basic Data Set; 2 methodological and statistical issue to avoid misinterpretation. *Spinal Cord*, 2010; **48**: 230–238.

- 2 Baunsgaard CB, Chhabra HS, Harvey LA, Savic G, Sisto SA, Qureshi F *et al.* Reliability of the International Spinal Cord Injury Musculoskeletal Basic Data Set. *Spinal Cord*; e-pub ahead of print 3 May 2016; doi:10.1038/sc.2016.42.
- 3 Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; **33**: 613–619.
- 4 Biering-Sørensen F, Burns AS, Curt A, Harvey LA, Jane Mulcahey M, Nance PW *et al.* International spinal cord injury musculoskeletal basic data set. *Spinal Cord* 2012; **50**: 797–802.
- 5 Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005; **85**: 257–268.