Methodologic Issues on Interrater Reliability Regarding Structural and DTI-Based Corticospinal Tract Asymmetry

F. Jalalvandi and M. Naderi

doi: https://doi.org/10.3174/ajnr.A5799
http://www.ajnr.org/content/39/11/E123
Methodologic Issues on Interrater Reliability Regarding Structural and DTI-Based Corticospinal Tract Asymmetry

We read, with interest, the article by Foesleitner et al1 published in the June 2018 issue of the American Journal of Neuroradiology. The purpose was to investigate a clinically feasible imaging approach to assess corticospinal tract (CST) asymmetry in unilateral polymicrogyria (PMG), check diffusion-based tractography as a guide to the presumed motor area within the dysplastic cortex, and investigate whether the “rule” of CST asymmetry as a good prognostic factor for postsurgical motor function preservation specifically applies to cases of unilateral PMG involving the central region. The interrater reliability was calculated by the Cohen coefficient.1 The authors reported that the interrater reliability in the assessment of corticospinal tract asymmetry was most robust at the level of the cerebral crus. Also, excellent congruence was reached by categorizing the asymmetry degree into no or minimal asymmetry or moderate/severe asymmetry (κ = 1.0) in the event that the other levels of assessment did not result in considerable agreement (κ = 0.21–0.6).1

It is of crucial importance to know that the κ value cannot be a sign of good agreement. In assessing the agreement of a qualitative variable, the κ value has 2 major weaknesses: 1) It depends on the prevalence in each class—that is, there might be different κ values of the same percentages for concordant and discordant cells. As can be seen in the Table, the prevalence of concordant cells in both (a) and (b) situations is 90%, while that of discordant cells is 10%. However, we can get different values of κ (0.44 and 0.80) for concordant and discordant cells, respectively. 2) The κ value also depends on the number of classes. It is preferable to use a weighted κ in such situations to obtain unbiased results.2–4

They concluded that visual assessment of structural and diffusion tensor images of the corticospinal tract (especially at the cerebral crus) is a reliable and clinically feasible imaging approach in the preoperative work-up of patients with unilateral PMG affecting the central region. Also, in noncompliant patients, DTI-based tractography is a useful alternative to task-based fMRI and helps in the anatomic localization of the primary motor cortex. If one considered the above-mentioned limitations of the κ value to assess reliability, such a conclusion may be misleading. Therefore, misinterpretation cannot be avoided.2–4

In this letter, we discuss the limitations of the κ value to assess reliability. Therefore, any conclusion especially in reliability analysis should be supported by the above mentioned statistical and methodological issues.

REFERENCES
2. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics. 2nd ed. Sudbury: Jones and Bartlett Publisher; 2007

F. Jalalvandi
M. Naderi
Department of Operating Room
School of Paramedical
Kermanshah University of Medical Sciences
Kermanshah, I.R. Iran

http://dx.doi.org/10.3174/ajnr.A5799