Letter to the Editor

Computed tomography assessment of anterior ethmoidal canal dehiscence: methodological issue on interobserver agreement

Dear Editor,

With deep interest we read the article by Guarnizo and colleagues, which was published in the April 2020 issue of the *Neuroradiology Journal*.[1] The aim of the authors was to investigate the interobserver agreement in the computed tomography (CT) assessment of the anterior ethmoidal canal (AEC) dehiscence.[1] In 199 patients, the presence of AEC dehiscence, the presence of paranasal sinuses (PNS) opacification and the best CT plane of AEC evaluation were assessed by two neuroradiologists separately. The kappa coefficient was used to evaluate the interobserver agreement. Based on the authors' results, the kappa coefficient for the interobserver agreement identification of AEC dehiscence, the AEC dehiscence in cases with opacification of ethmoidal air cells, the best planes for AEC assessment were $k = 0.24$, $k = 0.75$ and $k = 0.09$, respectively.[1]

After reading this article carefully and critically, we have some statistical and methodological concerns to point out, as follow: measuring the reliability of qualitative variables by kappa values has some limitations that we will mention below. First, the kappa value depends on the prevalence in each category. Second, it also depends on the number of categories.[2–6] We should mention that when a variable with more than two categories or an ordinal scale is used (with three or more ordered categories), then the weighted kappa would be a good choice. Table 1 shows the agreement by applying kappa (0.43 as moderate) and weighted kappa (0.63 as good) which has different values and consequently different interpretations.[2–6]

The authors concluded that the suboptimal interobserver agreement could potentially limit the usefulness of CT scans for routine assessment of AEC dehiscence also in patients with PNS opacification. CT scans could still add valuable information regarding AEC dehiscence.[1]

In this letter, we discuss the correct reporting of reliability and important limitations of applying kappa coefficient to assess reliability.[2–6] Any conclusion in reliability analyses needs to be supported by the methodological and statistical issues mentioned above. Otherwise, misinterpretation cannot be avoided.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Siamak Sabour https://orcid.org/0000-0002-1928-992X

Table 1. The kappa and weighted kappa values for calculating agreement between two neuroradiologists for more than two categories and depend on prevalence.

Anterior ethmoidal canal (AEC) dehiscence	Neuroradiologist 1			
	Yes	No	Unknown	Sum
Neuroradiologist 2				
Yes	60	20	1	81
No	2	12	4	18
Unknown	3	11	11	25
Sum	65	43	16	124
	Estimate			
Kappa	0.43			
Weighted kappa	0.63			

Cells indicate authors' own hypothetical numbers to compare the values of kappa and weighted kappa.

References

1. Guarnizo A, Nguyen TB, Glikstein R, et al. Computed tomography assessment of anterior ethmoidal canal dehiscence: an interobserver agreement study and review of the literature. *Neuroradiol J* 2020; 33: 145–151.
2. Szklo M and Nieto FJ. *Epidemiology Beyond the Basics*, 3rd ed. Manhattan, New York: Jones and Bartlett Publisher, 2014.
3. Naderi M and Sabour S. Reproducibility of diagnostic criteria associated with atypical breast cytology: a methodological issue. *Cytopathology* 2018; 29: 396.
4. Sabour S. Reproducibility of semi-automatic coronary plaque quantification in coronary CT angiography with sub-mSv radiation dose; common mistakes. *J Cardiovasc Comput Tomog*. 2016; 10: 21–22.

5. Naderi M and Sabour S. Inter and intraobserver reliability and critical analysis of the FFP classification of osteoporotic pelvic ring injuries: methodological issue. *Injury* 2019; 50: 1261–1262.

6. Sabour S. Reproducibility of the external surface position in left-breast DIBH radiotherapy with spirometer-based monitoring: methodological mistake. *J Appl Clin Med Phys* 2014; 15: 4909.

Mehdi Naderi[1] **and Siamak Sabour**[2,3,] ⓘ
[1]Clinical Research Development Centre, Kermanshah University of Medical Sciences, Iran

[2]Department of Clinical Epidemiology, Shahid Beheshti University of Medical Sciences, Iran
[3]Safety Promotions and Injury Prevention Research Centre, Shahid Beheshti University of Medical Sciences, Iran

Corresponding author:
Siamak Sabour, Department of Clinical Epidemiology, School of Health and Safety, Shahid Beheshti University of Medical Sciences, Chamran Highway, Velenjak, Daneshjoo Blvd, Tehran, IR Iran.
Email: s.sabour@sbmu.ac.ir