

Methodological and Statistical Issues in Diagnostic Research

Siamak Sabour^{1,2*}

¹Department of Clinical Epidemiology, School of Health, Shahid Beheshti University of Medical Sciences, Tehran, IR, Iran

²Safety Promotions and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran, IR, Iran

*Corresponding author: Siamak Sabour, Department of Clinical Epidemiology, School of Health, Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran, IR, Chamran Highway, Velenjak, Daneshjoo Blvd, Iran, Tel: +98-21- 22421814; E-mail: s.sabour@sbmu.ac.ir

Received date: September 09, 2017; Accepted date: September 15, 2017; Published date: September 19, 2017

Copyright: © 2017 Sabour S. This is an open-access article distributed under the terms of the creative commons attribution license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Editorial

Diagnostic researches are among interesting field of clinical researches. However, methodological and statistical issues in such researches are not being considered appropriately. Diagnostic value should be considered as diagnostic accuracy (validity) and diagnostic precision (reliability or agreement). In case of binary variable, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio positive (LR+), likelihood ratio negative (LR-) as well as odds ratio (ratio of true to false results) are the most appropriate estimates to evaluate validity of a test compared to a gold standard. Therefore, it is better to report all these validity estimates together. Otherwise, final interpretation will be confusing. Moreover, it is important to know that for clinical purposes, reporting diagnostic added value should be considered using receiver operating characteristic (ROC) curve because all the above mentioned validity estimates can be acceptable while diagnostic added value may be clinically negligible. Regarding quantitative variables, Interclass correlation coefficient (Pearson r or spearman rho) can be considered as an appropriate statistical test to assess validity [1-6].

Reliability (precision or agreement) as a different methodological issue of the diagnostic value should also be assessed using appropriate estimate. For qualitative variables, weighted kappa should be used with caution. Two important weaknesses of Cohen's kappa to assess agreement of a qualitative variable are as follows. First, it depends on the prevalence in each category, which means it can be possible to have different k values having the same percentage for both concordant and discordant cells. Table 1 shows that in both (a) and (b) situations, the prevalence of concordant cells are 90% and of discordant cells, 10%; however, we get different kappa values (0.44 as moderate and 0.80 as very good, respectively). Kappa value also depends on the number of categories. In such a situation, a weighted kappa is a preferable test, giving an unbiased result. Finally, the P value or 95% CI is not reported for a weighted kappa in reliability analysis, because statistically significant does not necessarily means clinically important. Regarding quantitative variables, Intraclass correlation coefficient (ICCC) agreement single measure and Bland Altman plot can be considered as appropriate tests to assess reliability [7-11].

Situation a		Observer 1		
		Prevalence of Positive Results (%)	Prevalence of Negative Results (%)	Total (%)
Observer 2	Prevalence of Positive Results	85	5	90
	Prevalence of Negative Results	5	5	10
	Total	90	10	100
k=0.44				
Situation b		Observer 1		
		Prevalence of Positive Results (%)	Prevalence of Negative Results (%)	Total (%)
Observer 2	Prevalence of Positive Results	45	5	50
	Prevalence of Negative Results	5	45	50
	Total	50	50	100
k=0.80				

Table 1: Comparison of Two Observers' Diagnoses with Different Prevalence in the Two Categories^a. Author's own hypothetical data to demonstrate the limitation of the kappa value to assess agreement.

In this editorial, I discuss the methodological and statistical issues in diagnostic researches. Therefore, conclusion of a diagnostic research should be supported by the above mentioned statistical and

methodological issues. Otherwise, misdiagnosis and mismanagement of the patients cannot be avoided.

References

1. Szklo M, Nieto FJ (2007) *Epidemiology beyond the basics*, 2nd (edn) Manhattan, Jones and Bartlett Publisher, New York, United States.
2. Sabour S, Ghassemi F (2016) Accuracy and reproducibility of the ETDRS visual acuity chart: methodological issues. *Graefes Arch Clin Exp Ophthalmol* 254: 2073-2074.
3. Sabour S (2015) Validity and reliability of the new Canadian Nutrition Screening Tool in the 'real-world' hospital setting: Methodological issues. *Eur J Clin Nutr* 69: 864.
4. Sabour S (2015) Spinal instability neoplastic scale: methodologic issues to avoid misinterpretation. *AJR Am J Roentgenol* 204: W493.
5. Sabour S (2014) Validity and reliability of the 13C-methionine breath test for the detection of moderate hyperhomocysteinemia in Mexican adults; statistical issues in validity and reliability analysis. *Clin Chem Lab Med* 52: e295-e296.
6. Sabour S (2013) A quantitative assessment of the accuracy and reliability of O-arm images for deep brain stimulation surgery. *Neurosurgery* 72: E696.
7. Sabour S (2016) Adherence to guidelines strongly improves reproducibility of brachial artery flow-mediated dilation. Common mistakes and methodological issue. *Atherosclerosis* 251: 490-491.
8. Sabour S (2016) Reliability of a new modified tear breakup time method: methodological and statistical issues. *Graefes Arch Clin Exp Ophthalmol* 254: 595-596.
9. Sabour S (2015) Reproducibility of dynamic Scheimpflug-based pneumotonometer and its correlation with a dynamic bidirectional pneumotonometry device: methodological issues. *Cornea* 34: e14-15.
10. Sabour S (2015) Reliability of automatic vibratory equipment for ultrasonic strain measurement of the median nerve: common mistake. *Ultrasound Med Biol* 41: 1119-1120.
11. Sabour S, Dastjerdi EV (2013) Reliability of four different computerized cephalometric analysis programs: a methodological error. *Eur J Orthod* 35: 848.