Contents lists available at ScienceDirect

Annals of Diagnostic Pathology

journal homepage: www.elsevier.com/locate/anndiagpath

Letter to the Editors

HER2 immunohistochemistry in 205 cases of invasive breast carcinoma additionally tested by ISH; Statistical issue on reproducibility to avoid misinterpretation

I was interested to read the papers by Casterá C and colleagues published in Apr 2020 issue of Ann Diagn Pathol [1]. Assessment of HER2 biomarker in invasive breast carcinoma patients allows a specific therapeutic approach. Clinical guidelines indicate immunohistochemistry (IHC) and in situ hybridization (ISH) to test HER2, however both have drawbacks which result in low reproducibility of results especially in equivocal cases. The authors aimed to quantify inter-observer IHC reproducibility and cross it with the ISH result. They sent 205 invasive breast carcinoma cases for ISH retest from 14 hospitals, 5 observers to assess the IHC and 2 observers for the ISH of each case. They mentioned that the observers only achieve an absolute agreement for IHC in 1 out of 3 cases. The inter-observer concordance for IHC is low ($0.2 \le k \le 0.4$) or moderate ($0.41 \le k \le 0.6$). In ISH positive cases the concordance for IHC is higher than in the ISH negative cases.

I want to congratulate the authors for this article, and make some contributions. The main purpose of my letter is to mention methodological limitations of kappa to assess agreement [2]. First, kappa depends on the prevalence in each category [2-6]. It is possible to have the prevalence of concordant cells equal to 90% and discordant cells to 10%; however, get different kappa value [0.44 as moderate vs. 0.81 as very good], respectively (Table 1). Kappa value also depends on the number of categories. [2-9]. As in this study, the possible diagnoses for both techniques, following the 2013 guidelines recommendations, are 4: Indeterminate, Negative, Equivocal and Positive. I should mention that applying the weighted kappa would be a good choice to assess intra-rater agreement (Table 2). However, Fliess kappa is suggested to assess inter-rater agreement when we have more than two raters. As in this study, the 5 IHC assessments of each case were made independently by 4 observers (O1, O2, O3 and O4) from the central laboratory and 1 external observer (O5) corresponding to the observer from the local laboratory requesting the ISH reflex test and sender of the IHC slide. They concluded that low and moderate IHC inter-observer concordance, finding the more worrying values among the ISH negative cases. Subjective interpretation of the techniques, among other factors, has negative impact in HER2 evaluation. Any conclusion on agreement should take into account the above-mentioned statistical issues and limitations of kappa. Otherwise, misinterpretation may occur.

Table 1
Limitation of Kappa to assess agreement between two observers with different prevalence in the two categories.

Observer 2		Negative	Observer 1 positive	Total (%)
Situation (a)	Negative	**85**	5	90
	Positive	5	**5**	10
K = 0.44 (moderate)	Total	90	10	100
Situation (b)	Negative	**45**	5	50
	Positive	5	**45**	50
K = 0.81 (very good)	Total	50	50	100

Bold indicates frequency of agreement cells.

Table 2
The kappa and weighted kappa values for calculating reproducibility between 2 reviewers for a variable with more than 2 categories.

	ISH/IHC	Observer 1			SUM
		Indeterminate	Negative	Equivocal	
Observer 2	Indeterminate	60	20	1	81
	Negative	2	12	4	18
	Equivocal	3	11	11	25
SUM		65	43	16	124
		Estimate			
KAPPA		0.43 (Moderate)			
WEIGHTED KAPPA		0.63 (Good)			

Immunohistochemistry (IHC).
In situ hybridization (ISH).

Source(s) of support

None.

Declaration of competing interest

None.

References

[1] Casterá C, Bernet L. HER2 immunohistochemistry inter-observer reproducibility in 205 cases of invasive breast carcinoma additionally tested by ISH. Ann Diagn Pathol 2020 Apr;45:151451. https://doi.org/10.1016/j.anndiagpath.2019.151451.

[2] Szklo M, Nieto FJ. Epidemiology beyond the basics. 3 rd ed. Manhattan, New York: Jones and Bartlett Publisher, United State; 2014. p. 313–43.

[3] 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(4):4909. https://doi.org/10.1120/jacmp. v15i4.4909.

[4] Sabour S. Reproducibility of semi-automatic coronary plaque quantification in coronary CT angiography with sub-mSv radiation dose; common mistakes. J Cardiovasc Comput Tomogr 2016;10(5):e21–2. https://doi.org/10.1016/j.jcct.2016.07.002.

[5] Sabour S, Ghassemi F. Accuracy, validity, and reliability of the infrared optical head tracker (IOHT). Invest Ophthalmol Vis Sci 2012;53(8):4776. https://doi.org/10.1167/iovs.12-10324.

[6] Naderi M, Sabour S. Reproducibility of diagnostic criteria associated with atypical breast cytology: a methodological issue. Cytopathology 2018;29(4):396. https://doi.org/10.1111/cyt.12560. [Epub 2018 May 23].

[7] Sabour S. Reproducibility of endometrial cytology by the Osaki Study Group Method: methodological issues. Cytopathology 2017;28(5):441–2. https://doi.org/10.1111/cyt.12447.

[8] Sabour S. Re: does the experience level of the radiologist, assessment in consensus, or the addition of the abduction and external rotation view improve the diagnostic reproducibility and accuracy of MRA of the shoulder? Clin Radiol 2015;70(3):333–4. https://doi.org/10.1016/j.crad.2014.11.011.

[9] Sabour S. A quantitative assessment of the accuracy and reliability of O-arm images for deep brain stimulation surgery. Neurosurgery 2013;72(4):E696. https://doi.org/10.1227/NEU.0b013e318282d66e.

Siamak Sabour[a,b,*]

[a] *Department of Clinical Epidemiology, School of Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

[b] *Safety Promotions and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

E-mail address: s.sabour@sbmu.ac.ir.

[*] Corresponding author at: Chamran Highway, Velenjak, Daneshjoo Blvd, Department of Clinical Epidemiology, School of Health, Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran 198353-5511, Iran.