

SHORT COMMUNICATION

Reliability and Validity Analysis in the Field of Tumor; Common Mistakes

Siamak Sabour

Siamak Sabour, Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran

Siamak Sabour, Department of Clinical Epidemiology, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran

Correspondence to: Siamak Sabour, MD, MSc, DSc, PhD, Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran.

Email: s.sabour@sbmu.ac.ir

Telephone: +98-21- 22421814

Received: January 16, 2014 Revised: February 5, 2014

Accepted: February 10, 2014

Published online: April 18, 2014

ABSTRACT

Reliability (precision) and validity (accuracy) are two important methodological issues in all fields of researches. The reliability is being assessed by inappropriate tests which all of them are among common mistakes. For quantitative variable Intra Class Correlation Coefficient (ICC) and for qualitative variables weighted kappa should be used. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio positive and likelihood ratio negative as well as diagnostic accuracy and odds ratio are among the tests to evaluate the validity of a single test compared to a gold standard. For reliability and validity analysis, appropriate tests should be applied by clinical researchers. Otherwise, misdiagnosis and mismanagement of the patients in routine clinical care cannot be avoided using inappropriate tests to assess reliability and validity.

© 2014 ACT. All rights reserved.

Key words: Reliability, Validity, Diagnostic tests, Mistake

Sabour S. Reliability and Validity Analysis in the Field of Tumor; Common Mistakes. *Journal of Tumor* 2014; 2(4): 122-124 Available from: URL: <http://www.ghrnet.org/index.php/JT/article/view/652>

ADVANCES IN KNOWLEDGE

1. Reliability (precision) and validity (accuracy) are two important

methodological issues in all fields of researches.

2. The reliability is being assessed by inappropriate tests which all of them are among common mistakes and is being published by high impact journals.

3. As a take home message, for reliability and validity analysis, appropriate tests should be applied by clinical researchers.

IMPLICATION FOR PATIENT CARE

Misdiagnosis and mismanagement of the patients in routine clinical care cannot be avoided using inappropriate tests to assess reliability and validity.

RELIABILITY AND VALIDITY ANALYSIS IN THE FIELD OF TUMOR; COMMON MISTAKES

Reliability (precision) and validity (accuracy) are two completely different and important methodological issues in all fields of researches. Reliability (repeatability or reproducibility) is being assessed by different statistical tests such as Pearson, least square and paired *t test* which all of them are among common mistakes in reliability analysis (Figure 1, 2)^[1] and is being published by high impact journals^[2-8].

Briefly, for quantitative variable Intra Class Correlation Coefficient (ICC) and for qualitative variables weighted kappa should be used with caution because kappa has its own limitation too^[1,9-40].

It is crucial to know that there is no value of kappa that can be regarded universally as indication good agreement. Two important weaknesses of kappa value to assess agreement of a qualitative variable are as follow: It depends upon the prevalence in each category which means it can be possible to have different kappa value having the same percentage for both concordant and discordant cells! Figure 3 shows that in both (a) and (b) situations the prevalence of concordant cells are 80% and discordant cells are 20%, however, we get different kappa value (0.38 and 0.60) respectively. Kappa value also depends upon the number of categories which means the higher the categories, the lower the amount of kappa value^[9-40].

Sensitivity [Percent with the disease who test positive, True Positives / (True Positives + False Negative)], specificity [Percent healthy who test negative, True Negatives / (True Negatives + False

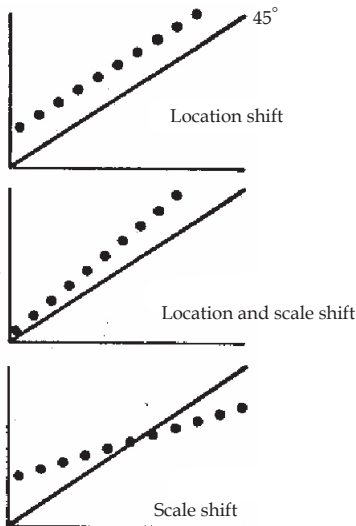


Figure 1 Cases when Pearson correlation coefficient fails to detect non reproducibility.

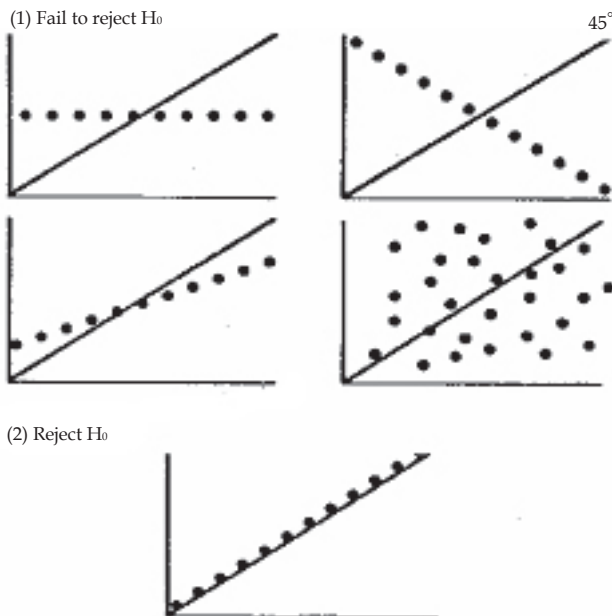


Figure 2 Cases when paired t-test can be misleading. H₀: Means are equal.

(a)

		Observer 1		Total
		Positive	Negative	
Observer 2	Positive	70	10	80
	Negative	10	10	20
	Total	80	20	100

K=0.38

(b)

		Observer 1		Total
		Positive	Negative	
Observer 2	Positive	40	10	50
	Negative	10	40	50
	Total	50	50	100

K=0.60

Figure 3 Comparison of two observers' diagnosis with different prevalence in the two categories.

Positive)] positive predictive value (PPV), [Percent of positive tests who actually are diseased, True Positives / (True Positives + False Positive)], negative predictive value (NPV) [Percent of negative tests who are healthy, True Negatives / (True Negatives + False

Negative)], likelihood ratio positive and likelihood ratio negative as well as diagnostic accuracy [(both true positive and true negative results / total)* 100] and odds ratio (true results / false results) preferably more than 50, are among the tests to evaluate the validity of a single test compared to a gold standard^[9-40].

As a take home message, for reliability and validity analysis, appropriate tests should be applied by researchers. Otherwise, misdiagnosis and mismanagement of the patients cannot be avoided.

CONFLICT OF INTERESTS

There are no conflicts of interest with regard to the present study.

REFERENCES

- Lawrence I, Kuei Lin. A Concordance Correlation Coefficient to Evaluate Reproducibility. *BIOMETRICS* 1989 March; 45: 255-268
- Jain V, Duda J, Avants B, Giannetta M, Xie SX, Roberts T, Detre JA, Hurt H, Wehrli FW, Wang DJ. Longitudinal Reproducibility and Accuracy of Pseudo-Continuous Arterial Spin-labeled Perfusion MR Imaging in Typically Developing Children. *Radiology* 2012 May; 263(2): 527-536
- Albarakati SF, Kula KS, Ghoneima AA, The reliability and reproducibility of cephalometric measurements: a comparison of conventional and digital methods. *Dentomaxillofac Radiol* 2012 Jan; 41(1): 11-17
- Lai V, Tsang WK, Chan WC, Yeung TW. Diagnostic accuracy of mediastinal width measurement on posteroanterior and anteroposterior chest radiographs in the depiction of acute nontraumatic thoracic aortic dissection. *Emerg Radiol* 2012 Mar 14.
- Ling LF, Obuchowski NA, Rodriguez L, Popovic Z, Kwon D, Marwick TH. Accuracy and Interobserver Concordance of Echocardiographic Assessment of Right Ventricular Size and Systolic Function: A Quality Control Exercise. *J Am Soc Echocardiogr* 2012 Apr 26
- Sato T, Tsujino I, Ohira H, Oyama-Manabe N, Yamada A, Ito YM, Goto C, Watanabe T, Sakaue S, Nishimura M. Validation study on the accuracy of echocardiographic measurements of right ventricular systolic function in pulmonary hypertension. *J Am Soc Echocardiogr* 2012 Mar; 25(3): 280-286
- Maislin G, Ahmed MM, Gooneratne N, Thorne-Fitzgerald M, Kim C, Teff K, Arnardottir ES, Benediktsdottir B, Einarsson H, Juliusson S, Pack AL, Gislason T, Schwab RJ. Single Slice vs. Volumetric MR Assessment of Visceral Adipose Tissue: Reliability and Validity Among the Overweight and Obese. *Obesity (Silver Spring)* 2012 Mar 7. doi: 10.1038/oby.2012.53
- Soviero VM, Leal SC, Silva RC, Azevedo RB. Validity of MicroCT for in vitro detection of proximal carious lesions in primary molars. *J Dent* 2012 Jan; 40(1): 35-40
- Jeckel J.F, Katz. D.L, Elmore, J.G, Wild, D.M.G, Epidemiology, Biostatistics and Preventive Medicine, 3rd edition. 2007, SAUNDERS, Elsevier, Philadelphia, PA, United State
- Kenneth J. Rothman, Sander Greenland, Timothy L. Lash. Modern Epidemiology, 4th edition. 2010. Lippincott Williams & Wilkins, Baltimore, United States
- Szklo M, Nieto. F.J, Epidemiology beyond the basics, 2nd edition, 2007, Jones and Bartlett Publisher, Manhattan, new York, United State
- Sabour S, Dastjerdi EV. Reliability of four different computerized cephalometric analysis programs: a methodological error. *Eur J Orthod* 2013 Dec; 35(6): 848. doi: 10.1093/ejo/cjs074. Epub 2013 Oct 16.
- Sabour S. Reliability and accuracy of skeletal muscle imag-

- ing in limb-girdle muscular dystrophies. *Neurology* 2013 Jun 11; **80(24)**: 2275. doi: 10.1212/WNL.0b013e318299ef6b
- 14 Sabour S, Ghassemi F. Predictive value of confocal scanning laser for the onset of visual field loss. *Ophthalmology* 2013 Jun; **120(6)**: e31-2. doi: 10.1016/j.ophtha.2013.01.055.
 - 15 Sabour S, Kermani H. A comparison between dental measurements taken from CBCT models and those taken from a digital method: a query about methodology. *Eur J Orthod* 2013 Oct; **35(5)**: 714-715. doi: 10.1093/ejo/cjt022. Epub 2013 May 29
 - 16 Sabour S. Interlaboratory and interstudy reproducibility of a novel lateral-flow device: a statistical issue. *J Clin Microbiol* 2013 May; **51(5)**: 1652. doi: 10.1128/JCM.00111-13
 - 17 Sabour S, Ghassemi F. Interrater reliability of intensive care unit electroencephalogram revised terminology: pitfalls and challenges of using kappa value. *J Clin Neurophysiol* 2013 Apr; **30(2)**: 210. doi: 10.1097/WNP.0b013e31827edcca.
 - 18 Sabour S. A quantitative assessment of the accuracy and reliability of O-arm images for deep brain stimulation surgery. *Neurosurgery* 2013 Apr; **72(4)**: E696. doi: 10.1227/NEU.0b013e318282d66e.
 - 19 Sabour S. Single slice vs. volumetric MR assessment of visceral adipose tissue: reliability and validity among the overweight and obese. *Obesity (Silver Spring)* 2013 Jan; **21(1)**: 6-7. doi: 10.1002/oby.20093.
 - 20 Sabour S, Ghorbani Z. Developmental fluoride neurotoxicity: clinical importance versus statistical significance. *Environ Health Perspect* 2013 Mar; **121(3)**: A70. doi: 10.1289/ehp.1206192.
 - 21 Sabour S. Effect of Training Status on Reliability of 1RM Testing in Women, Methodological Mistake. *J Strength Cond Res* 2013 Feb 25. [Epub ahead of print]
 - 22 Sabour S, Dastjerdi EV, Moezizadeh M. Accuracy of peri-implant bone thickness and validity of assessing bone augmentation material using cone beam computed tomography--is this correct? *Clin Oral Investig* 2013 Sep; **17(7)**: 1785. doi: 10.1007/s00784-013-0944-0. Epub 2013 Feb 21.
 - 23 Accuracy of linear intraoral measurements using cone beam CT and multidetector CT: methodological mistake. Sabour S, Kermani H, Dentomaxillofac Radiol. 2013;42(4):20130048. doi: 10.1259/dmfr.20130048. Epub 2013 Feb 18. No abstract available.
 - 24 Sabour S, Moezizadeh M, Dastjerdi EV. Reliability of shade selection using an intraoral spectrophotometer: common mistakes in reliability analysis. *Clin Oral Investig* 2013 Apr; **17(3)**:1025. doi: 10.1007/s00784-013-0930-6. Epub 2013 Feb 13.
 - 25 Sabour S, Vahid-Dastjerdi E, Validity of the bolton index using cone-beam computed tomography (CBCT); methodological mistake. *Med Oral Patol Oral Cir Bucal* 2013 Sep 1; **18(5)**:e822-3
 - 26 Sabour S, Dastjerdi EV. Validity of a visual scoring method for masticatory ability using test gummy jelly: methodological mistake. *Gerodontology* 2013 Mar; **30(1)**: 85. doi: 10.1111/ger.12007.
 - 27 Sabour S, Dastjerdi EV. Reliability of assessment of nasal flow rate for nostril selection during nasotracheal intubation: common mistakes in reliability analysis. *J Clin Anesth* 2013 Mar; **25(2)**:162. doi: 10.1016/j.jclinane.2012.10.006. Epub 2013 Jan 16.
 - 28 Sabour S, Dastjerdi EV. Reliability of implant surgical guides based on soft-tissue models: a methodological mistake. *J Oral Implantol* 2012 Dec; **38(6)**: 805. doi: 10.1563/AAID-JOID-12-00176.
 - 29 Sabour S, Ghassemi F, Ocular duction studies: statistical issues. *Ophthalmology* 2013 Jan; **120(1)**: 222-223. doi:10.1016/j.ophtha.2012.08.020.
 - 30 Sabour S. Recovery, dependence or death after discharge. *J Gen Intern Med* 2013 Mar; **28(3)**: 342. doi: 10.1007/s11606-012-2311-2.
 - 31 Sabour S, Dastjerdi EV, Moezizadeh M, Geha H. Effect of JPEG compression on the diagnostic accuracy of periapical images; mistakes and misinterpretations. *Dent Traumatol* 2013 Jun; **29(3)**: 251. doi: 10.1111/edt.12014. Epub 2012 Nov 6
 - 32 Sabour S. Reliability and repeatability of toe pressures measured with laser Doppler and portable and stationary photoplethysmography devices. *Ann Vasc Surg* 2012 Nov; **26(8)**: 1167. doi: 10.1016/j.avsg.2012.05.008.
 - 33 Sabour S. Validity of chronic obstructive pulmonary disease diagnoses in a large administrative database: The rule of thumb in validity analysis of a test. *Can Respir J* 2012 Sep-Oct; **19(5)**:331; author reply 331.
 - 34 Sabour S. Commentary on "Reliability of two different pre-surgical preparation methods for implant dentistry based on panoramic radiography and cone-beam computed tomography in cadavers". *J Periodontal Implant Sci*, 2012 Aug; **42(4)**: 144. doi: 10.5051/jpis.2012.42.4.144. Epub 2012 Aug 31.
 - 35 Sabour S, Ghassemi F. The reproducibility of measurements of differential renal function in paediatric 99mTc-MAG3 renography: is this correct? *Nucl Med Commun* 2012 Dec; **33(12)**: 1311; author reply 1311-2. doi:10.1097/MNM.0b013e328359453a.
 - 36 Sabour S. Reliability of subjective, linear, ratio, and area cephalometric measurements in assessing adenoid hypertrophy among different age groups. *Angle Orthod* 2012 Sep; **82(5)**: 948; author reply 949-50. doi: 10.2319/0003-3219-82.5.948.
 - 37 Sabour S, Dastjerdi EV. Reliability of Soft Tissue Model Based Implant Surgical Guides; A Methodological Mistake. *J Oral Implantol* 2012 Aug 20. [Epub ahead of print]
 - 38 Sabour S, Ghassemi F. Accuracy, validity, and reliability of the infrared optical head tracker (IOHT). *Invest Ophthalmol Vis Sci* 2012 Jul 13; **53(8)**: 4776. doi: 10.1167/iovs.12-10324.
 - 39 Sabour S, Ghassemi F. Validity of self-reported eye disease and treatment: is that correct?! *Br J Ophthalmol* 2012 Oct; **96(10)**: 1359. Epub 2012 Jul 11
 - 40 Sabour S, Ghassemi F. Reliability and validity of conjunctival ultraviolet autofluorescence measurement. *Br J Ophthalmol* 2012 Sep; **96(9)**: 1271; author reply 1271. doi: 10.1136/bjophthalmol-2012-302087. Epub 2012 Jun 13

Peer reviewers: Siamak Sabour, MD, PhD, Community Oral Health, Clinical Epidemiology, School of Dentistry, Chamran highway, Tabnak Blv. Velenjac, Tehran, Iran; Siamak Sabour, MD, PhD, Community Oral Health, Clinical Epidemiology, School of Dentistry, Chamran highway, Tabnak Blv. Velenjac, Tehran, Iran.