

A Validation Study of Automated Bone Scan Index: Effect on Reproducibility Due to the Procedural Variability in Bone Scan Image Acquisition. A Common Mistake

Siamak Sabour, MD, MSc, DSc, PhD, Postdoc ^{1,2}

¹ Safety Promotion and Injury Prevention Research Center, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran

² Department of Clinical Epidemiology, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran

Corresponding contributor: Siamak Sabour

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

E: s.sabour@sbmu.ac.ir

T: +98-21- 22421814

Type of article: Letter

Total number of pages: 1

Total number of photographs: 0

Source(s) of support: None

Conflicts of interest: No

A Validation Study of Automated Bone Scan Index: Effect on Reproducibility Due to the Procedural Variability in Bone Scan Image Acquisition. A Common Mistake

I was interested to read the paper by Anand D and colleagues published in the Dec 2016 edition of the J Nucl Med.¹ The purpose of the authors was to assess the impact of the variability in scanning speed and in vendor-specific γ -camera on reproducibility and accuracy of the automated bone scan index (BSI).¹ They measured reproducibility as the absolute difference between the repeated BSI values, and accuracy as the absolute difference between the observed BSI and the phantom-BSI values. Descriptive statistics were used to compare the generated data.¹

Reproducibility (precision) and validity (accuracy) as two completely different methodological issues should be assessed using appropriate tests. It is crucial to know that, regarding reliability, for quantitative variable Intra Class Correlation Coefficient (ICC) and for qualitative variables weighted kappa should be used with caution. However to assess validity, for quantitative variables, interclass correlation coefficient (Pearson r) and for qualitative variables, 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 most appropriate tests. Moreover, for reliability analysis, an individual based approach should be applied using ICC agreement single measure because approaching a global average (absolute difference) can simply cause misleading messages. Absolute difference can be almost the same with no reliability at all and the other way around.²⁻⁸

Based on their results, in the patient study, 75 patients, 25 in each group, were enrolled. The reproducibility of Grp2 (mean \pm SD, 0.35 \pm 0.59) was observed to be significantly lower than that of Grp1 (mean \pm SD, 0.10 \pm 0.13; $P < 0.0001$) and that of Grp3 (mean \pm SD, 0.09 \pm 0.10; $P < 0.0001$). However, no significant difference was observed between the reproducibility of Grp3 and Grp1 ($P = 0.388$).¹ Statistically significant and clinically importance are two completely different issue and in clinical research especially in reliability analysis, we should not emphasize on significant level (P.value).²⁻⁸

They concluded that the automated BSI accuracy and reproducibility were dependent on scanning speed but not on the vendor-specific γ -cameras. Such conclusion should be supported by the above mentioned statistical and methodological issue. Otherwise, in clinical practice, misdiagnosis and mismanagement of the patients may occur.

KEYWORDS: bone scan; bone scan index; imaging biomarker; metastatic prostate cancer; pre-analytical validation

REFERENCES

1. Anand A, Morris MJ, Kaboteh R, Reza M, Trägårdh E, et al. A Preanalytic Validation Study of Automated Bone Scan Index: Effect on Accuracy and Reproducibility Due to the Procedural Variabilities in Bone Scan Image Acquisition. *J Nucl Med*. 2016 Dec;57(12):1865-1871. Epub 2016 Jul 21.
2. Szklo M, Nieto. F.J, *Epidemiology beyond the basics*, 2nd edition, Manhattan, new York, United State, Jones and Bartlett Publisher, 2007.
3. Sabour S. Myocardial blood flow quantification by Rb-82 cardiac PET/CT: Methodological issues on reproducibility study. *J Nucl Cardiol*. 2016 Sep 6. [Epub ahead of print]
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 Sep-Oct;10(5):e21-2. doi: 10.1016/j.jcct.2016.07.002. Epub 2016 Jul 9.
5. Sabour S. Reliability of a new modified tear breakup time method: methodological and statistical issues. *Graefes Arch Clin Exp Ophthalmol*. 2016 Mar;254(3):595-6. doi: 10.1007/s00417-015-3138-4. Epub 2015 Aug 28.
6. Sabour S, Farzaneh F, Peymani P. Evaluation of the sensitivity and reliability of primary rainbow trout hepatocyte vitellogenin expression as a screening assay for estrogen mimics: Methodological issues. *Aquat Toxicol*. 2015 Jul;164:175-6. doi: 10.1016/j.aquatox.2015.05.003. Epub 2015 May 5.
7. 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 Mar;70(3):333-4. doi: 10.1016/j.crad.2014.11.011. Epub 2015 Jan 1.
8. Sabour S. The reliability of routine clinical post-processing software in assessing potential diffusion-weighted MRI "biomarkers" in brain metastases, common mistake. *Magn Reson Imaging*. 2014 Nov;32(9):1162. doi: 10.1016/j.mri.2014.08.002. Epub 2014 Aug 8.