importance are two completely different issues, and in clinical research—especially reliability analysis—we should not put the emphasis on significance level (*P* value) (*2–8*).

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

REFERENCES

1. Anand A, Morris MJ, Kaboteh R, 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;57: 1865–1871.
2. Szklo M, Nieto FJ. *Epidemiology Beyond the Basics.* 2nd ed. Manhattan, NY: Jones and Bartlett; 2007.
3. Sabour S. Myocardial blood flow quantification by Rb-82 cardiac PET/CT: methodological issues on reproducibility study. *J Nucl Cardiol.* September 6, 2016 [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;10:e21–e22.
5. Sabour S. Reliability of a new modified tear breakup time method: methodological and statistical issues. *Graefes Arch Clin Exp Ophthalmol.* 2016;254:595–596.
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;164:175–176.
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? [comment]. *Clin Radiol.* 2015;70:333–334.
8. Sabour S. The reliability of routine clinical post-processing software in assessing potential diffusion-weighted MRI "biomarkers" in brain metastases, common mistake [comment]. *Magn Reson Imaging.* 2014;32:1162.

Siamak Sabour
Shahid Beheshti University of Medical Sciences
Tehran 1413663633, Islamic Republic of Iran
E-mail: s.sabour@sbmu.ac.ir

Published online Jan. 12, 2017.
DOI: 10.2967/jnumed.115.156745

REPLY: Thank you for your interest and comments regarding our work (*1*). We agree that reproducibility/reliability and accuracy/validity are two different methodologic issues and should be evaluated using appropriate tests for each. Further, we also concur that for quantitative variables, the intraclass correlation coefficient should be used to test reliability and the interclass correlation coefficient (Pearson *r*) should be used to test validity. Since the automated bone scan index (BSI) is a quantitative parameter, the comments on qualitative assessments are inapplicable.

Our analytic validation study demonstrated the validity of the automated BSI with a Pearson *r* of 0.99 ($P < 0.0001$) and with associated parameters of linear regression (slope, 0.80 [95% confidence interval, 0.78–0.83]; intercept, 0.38 [95% confidence interval, 0.25–0.51]) (*2*). In this study, Cohen κ-agreement and other well-known standard tests for imaging biomarkers (*3*) were also used to evaluate the reliability of the automated BSI platform. These tests demonstrated a coefficient of variation (<20%) with a consistent linearity—satisfying the homoscedasticity and Shapiro–Wilken test—from low-burden to high-burden disease. The Bland–Altman plot of absolute difference in BSI readings of repeated scans revealed the analytic noise with an SD of 0.15. The study concluded that in the clinical setting, the assessment of change in BSI above the noise threshold (2 SDs, or 0.30) had high interobserver agreement (Cohen $\kappa = 0.96$).

Despite the analytic validation of BSI, variation in the imaging procedure can significantly affect the analytic performance of the imaging biomarker. Therefore, in a subsequent publication, we analyzed the effect of variability in scanning speed and γ-camera settings on the BSI assessment (*1*). The objective of our preanalytic study was not to evaluate the performance characteristics of BSI but to use appropriate statistical tests to evaluate the effect of procedural variability on the analytically validated BSI values. The Wilcoxon/Mann–Whitney test (*4*) demonstrated that the noise in the BSI value was significantly higher because of variability in scanning speed in comparison with the known analytic gold standard (accuracy performance) and with repeated measurement of test–retest patient bone scans (reproducibility performance). The *P* value in our study, from the Wilcoxon/Mann–Whitney test, did not imply clinical significance but referred to the statistical significance of the degree of systematic difference between the two BSI readings of the repeated bone scans that were performed in a specific study design to assess the effect of procedural variability on the known performance characteristics of BSI.

In conclusion, the statistical methodology must be reviewed in the context of the study design and its objective. We have performed a comprehensive assessment to demonstrate the reliability and validity of BSI, and with the preanalytic study, we demonstrated that the BSI performance characteristics were dependent on scanning speed. The study added empiric evidence toward the standardization of bone scan acquisitions for robust quantitative BSI assessment in multiinstitutional studies. Together, the analytic and preanalytic studies served as the foundation for prospective clinical investigations aimed at validating automated BSI as a quantitative imaging biomarker.

REFERENCES

1. Anand A, Morris MJ, Kaboteh R, 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;57: 1865–1871.
2. Anand A, Morris MJ, Kaboteh R, et al. Analytic validation of the automated bone scan index as an imaging biomarker to standardize quantitative changes in bone scans of patients with metastatic prostate cancer. *J Nucl Med.* 2016;57:41–45.
3. O'Connor JP, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol.* 2017;14:169–186.
4. Szklo M, Nieto FJ. *Epidemiology Beyond the Basics.* 2nd ed. Manhattan, NY: Jones and Bartlett; 2007.

Aseem U. Anand*
David Minarik
**Lund University*
Waldenströms Gata 5, Second Floor
Malmö 20502, Sweden
E-mail: aseem.anand@med.lu.se

Published online Apr. 6, 2017.
DOI: 10.2967/jnumed.117.190645