Journal of Nuclear Medicine, published on April 6, 2017 as doi:10.2967/jnumed.117.190645

Reply to comments by Siamak Sabour

Aseem U. Anand[1], David Minarik[2]

[1]Department of Translational Medicine Division of Urological Cancer, Lund University, Malmö, Sweden.

[2]Department of Radiation Physics, Skåne University Hospital, Malmö, Sweden

Aseem Undvall Anand, PhD

Department of Translational Medicine Division of Urological Cancer, Lund University

Waldenströms gata 5, SE 205 02 Malmö, Sweden

E-mail: aseem.anand@med.lu.se

Phone: +46706604084; Fax: +4640336911

Reply to comments by Siamak Sabour

Dear Dr. Sabour,

Thank you for your interest and comments regarding our work [1]. We agree that reproducibility/reliability and accuracy/validity are two different methodological issues and should be evaluated using appropriate tests for each. Further, we also concur that Intra Class Correlation and weighted kappa test reliability, and interclass correlation coefficient (Pearson's r) test the validity of the quantitative variables. Since automated BSI is a quantitative parameter of bone scans, the comments on the qualitative assessments are inapplicable.

Our analytical validation study demonstrated the validity of the automated BSI with Pearson's r of 0.99 ($p<0.0001$) [2]. In this study, Cohen's Kappa agreement and other well-known standard tests for imaging-biomarkers [3], were also employed to evaluate the reliability of 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 repeat scans revealed the analytical "noise" with Standard Deviation (SD) of 0.15. The study concluded that in clinical setting the assessment of change in BSI above the "noise" threshold (2 x SD=0.30) had high inter-observer agreement (Cohen's Kappa=0.96).

Despite the analytical validation of BSI, variation in procedural factors of the imaging modality can have a significant impact on the analytical performance of the imaging biomarker. Therefore, in subsequent publication, we analyzed the effect of the variability in scanning speed and gamma cameras on the Bone Scan Index (BSI) assessment [1]. The objective of our pre-analytical 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 demonstrated that [4] the “noise” in BSI value, was significantly higher due to variability in scanning speed against the known analytical gold-standard (accuracy performance) and against repeated measurement of test/re-test 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 in the degree of systematic difference between the two BSI readings of the repeated bone scans that were performed in 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 comprehensive assessment to demonstrate the reliability and validity of BSI, and with the pre-analytical study, we demonstrated that the BSI performance characteristics were dependent on scanning speed of bone scan. The study added to the empirical evidence towards the standardization of bone scan image acquisition for robust quantitative BSI assessment in multi-institutional studies. Together, the analytical and the pre-analytical study served as the foundation for prospective clinical investigations aimed to validate automated BSI as a quantitative imaging-biomarker indicative of assessing change in bone scan that is clinically relevant.

References:

1. Anand, A., 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**(12): p. 1865-1871.
2. Anand, A., 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**(1): p. 41-5.
3. O'Connor, J.P., et al., *Imaging biomarker roadmap for cancer studies.* Nat Rev Clin Oncol, 2017. **14**(3): p. 169-186.
4. Szklo, M. and J. Nieto, *Epidemiology beyond the basics.* Jones and Bartlett Publisher, 2007. **2 nd edition**.