DEPARTMENTS
Letters to the Editor

Self-Reported Weight and Height in Nuclear Medicine Patients: A Common Mistake Confusing Reliability and Accuracy

TO THE EDITOR: I read with great interest the article by Blum et al. recently published in the *Journal of Nuclear Medicine Technology* (*1*). The authors aimed to assess the reliability of the self-reported weight and height of nuclear medicine patients in view of recommendations for weight-dependent tracer application for imaging and therapy. In total, 824 patients (334 men and 490 women) were asked to report their weight and height before imaging or therapy, along with their level of confidence that the weight and height they were reporting were correct. Subsequently, the weight and height of each patient were measured, and body mass index, body surface area, and lean body mass were calculated. Differences between the reported and true values were compared for statistical significance. The results indicated that an over- or underestimation of weight by at least 10% was observed in 2% of the patients, and height was overestimated by 1% of the patients. Surprisingly, the authors concluded that most self-reported weights and heights of nuclear medicine patients are accurate.

However, there were some methodologic issues regarding accuracy and reliability. First, it is crucial to realize that accuracy and reliability are two completely different methodologic issues. The term *accuracy* means the degree to which the result of a measurement, calculation, or specification conforms to the correct value or a standard. In other words, accuracy is the most important criterion for the quality of a test and refers to whether the test measures what it claims to measure. The core design for determining and measuring the accuracy of a test is a comparison between an index test and a reference standard by applying both on similar people who are suspected of having the target result of interest. The term *reliability* denotes refinement of a measurement, calculation, or specification, especially as represented by the number of digits given. Accuracy studies should report significant and comprehensive information together with the absolute number of true-positive, false-positive, false-negative, and true-negative results or should provide information that allows calculation of a minimum of one diagnostic performance indicator (i.e., sensitivity, specificity, predictive values, or likelihood ratio). Therefore, we recommend applying the most appropriate estimates to evaluate the accuracy of the self-reported weight and height. The Pearson r or the Spearman ρ can be applied to assess accuracy for quantitative variables. However, for qualitative (binary) variables, some of the well-known ways to assess accuracy include sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio (ranging from 1 to infinity; the higher the positive likelihood ratio, the more accurate the test), negative likelihood ratio (ranging from 0 to 1; the lower the negative likelihood ratio, the more accurate the test), diagnostic accuracy, and odds ratio (ratio of true results to false results) (*2–8*).

Second, what is critically important is reliability, which is conceptually different from accuracy. Consequently, our methodologic and statistical approach to assessing reliability should be different. Depending on the type of variable, appropriate estimates to assess reliability are completely different from those used to assess accuracy. For quantitative variables, we can apply either the intraclass correlation coefficient or Bland–Altman plots. For qualitative variables, we can apply the weighted κ or the Fleiss κ to assess intra- or interobserver reliability, respectively.

Thus, because of the inappropriate use of statistical tests (Student *t* test and ANOVA) for accuracy and reliability analyses, as well as misinterpretation of the results, there may be a high level of uncertainty about the conclusion of Blum et al. The evidence is insufficient to conclude that the self-reported weights and heights of nuclear medicine patients are accurate.

REFERENCES

1. Blum KS, Büsch N, Beyer T, et al. In patients we trust: reliability of self-reported weight and height in nuclear medicine patients. *J Nucl Med Technol.* 2019;47:133–136.
2. Sabour S. Reliability of immunocytochemistry and fluorescence in situ hybridization on fine-needle aspiration cytology samples of breast cancers: methodological issues. *Diagn Cytopathol.* 2016;44:1128–1129.
3. Sabour S. Reliability of automatic vibratory equipment for ultrasonic strain measurement of the median nerve: common mistake. *Ultrasound Med Biol.* 2015;41:1119–1120.
4. Sabour S, Dastjerdi EV. Reliability of four different computerized cephalometric analysis programs: a methodological error. *Eur J Orthod.* 2013;35:848.
5. Sabour S. Reliability of the ASA physical status scale in clinical practice: methodological issues. *Br J Anaesth.* 2015;114:162–163.
6. Sabour S, Ghassemi F. The validity and reliability of a signal impact assessment tool: statistical issue to avoid misinterpretation. *Pharmacoepidemiol Drug Saf.* 2016;25:1215–1216.
7. Sabour S. Validity and reliability of the new Canadian Nutrition Screening Tool in the 'real-world' hospital setting: methodological issues. *Eur J Clin Nutr.* 2015;69:864.
8. Sabour S. Methodologic concerns in reliability of noncalcified coronary artery plaque burden quantification [reply]. *AJR.* 2014;203:W343.

Siamak Sabour
Shahid Beheshti University of Medical Sciences
Tehran, Islamic Republic of Iran 198353-5511
E-mail: s.sabour@sbmu.ac.ir

Is There a Need for a Pediatric PET/CT Camera?

TO THE EDITOR: PET imaging is commonly used in adult patients. Its use in pediatric patients is limited but indicated in certain cancers such as lymphoma, sarcoma, neuroblastoma, and central nervous system tumors (*1*). The main limitation of PET imaging is its relatively limited spatial resolution in detecting small lesions, mainly because of the noncollinearity of the 511-keV annihilation photons, positron range, detector scatter, and parallax error (*2–5*). However, small-animal PET scanners have higher spatial resolution than standard human PET scanners (approximately 1 mm vs. 4–6 mm)