

Marker Selection Based on Only Reproducibility Can Be Questioned

TO THE EDITOR: An excellent recently published paper (1) suggests a role for image heterogeneity markers, with an eye toward therapy monitoring. The reproducibilities of 24 individual measuring processes were determined from esophageal tumor scans obtained a few days apart. The “Conclusion” section singles out 5 markers as being preferred, as justified solely from their being the most reproducible (i.e., the 5 with the lowest tabulated SD for percentage reproducibility).

Although readers can examine the text beyond the “Conclusion,” they might preferably consult a related interesting and thorough investigation published earlier by the authors (2) to see a fuller picture for choosing the best markers. This earlier work on selecting markers correctly used receiver-operating-characteristic (ROC) curves requiring outcome information that indicates therapy responders versus others. For the particular circumstances encountered here, this ROC-based selection methodology gave results similar to a solely reproducibility-based selection methodology. The “Conclusion” does not emphasize that only ¹⁸F-FDG PET esophageal studies were conducted.

The need to also have outcome information seems intuitive, as is mathematically supported by the exclusive analytic dependency (3) of the ROC area on a ratio (referred to as the Mahalanobis distance in classification theory and popularly also relatedly encountered in *t* testing between groups):

$$\frac{\text{difference in average marker values of 2 states}}{\left[(\text{measuring process SDs})^2 + (\text{states' SDs})^2 \right]^{1/2}}$$

Noteworthy besides the measurement reproducibility term are 2 biologically based terms: how much the markers of 2 states (e.g., for responders vs. others) differ on average, and what variability in markers is additionally encountered because of biologic randomness. It cannot be assumed that the relative importance among these 3 terms for all tumors, marker types, and protocols will resemble that encountered for the heterogeneity markers of esophageal tumors on ¹⁸F-FDG PET scans.

Finally, the traditional full-range ROC area, though quite good, may not be the entire story for marker selection in all instances. Practical clinical considerations may sometimes lead an investigator to conclude that only some judiciously chosen sensitivity or specificity range, defining a partial area, is of importance. Also, though complex to study, cost-to-benefit considerations can play a role when markers come from substantially differing protocols.

REFERENCES

1. Tixier F, Haat M, Le Rest CC, et al. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in ¹⁸F-FDG PET. *J Nucl Med.* 2012;53:693–700.
2. Tixier F, Chez Le Rest C, Haat M, et al. Intratumor heterogeneity characterized by textural features on baseline ¹⁸F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med.* 2011;52:369–378.

3. Swets JA, Pickett RM. *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory.* New York, NY: Academic Press; 1982.

Joseph A. Thie
University of Tennessee
 12334 Bluff Shore Dr.
 Knoxville, TN 37922
 E-mail: jathie@utk.edu

Published online Oct. 12, 2012.
 DOI: 10.2967/jnumed.112.108985

REPLY: We appreciate your attention and interest regarding our work in the field of new PET-image-derived indices within the context of therapy response prediction and prognosis. Our recent article (1) was concerned with the physiologic reproducibility of such tumor-derived indices based on uptake heterogeneity features. The objective of the work was to provide support for the eventual introduction of these heterogeneity parameters into clinical use. The physiologic reproducibility of any image-derived index is indeed a necessary condition, as is a robust methodology for its calculation, along with standardization in terms of reconstruction and associated correction methodologies. Satisfying these conditions reduces the variance of such parameters and the associated reproducibility limits, allowing smaller parameter changes during treatment to become significant. This in turn can lead to response assessment earlier during the course of treatment and going forward, facilitating the prediction of response directly from baseline PET images (2). However, although a prerequisite, none of these conditions alone is sufficient or the only determinant in ensuring that any PET-image-derived parameter has an eventual impact within the context of therapy response. In addition to clinical research studies already performed (2,3), studies using the tumor heterogeneity parameters showing the best physiologic reproducibility and robustness are needed for different cancer models and treatment regimes to demonstrate whether these new PET-image-derived parameters will be of interest in clinical practice.

REFERENCES

1. Tixier F, Haat M, Cheze Le Rest C, et al. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in ¹⁸F-FDG PET. *J Nucl Med.* 2012;53:693–700.
2. Tixier F, Cheze Le Rest C, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline ¹⁸F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med.* 2011;52:369–378.
3. El Naqa I, Grigsby PW, Aptea A, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit.* 2009;42:1162–1171.

Florent Tixier*
Dimitris Visvikis
Mathieu Hatt
Laurent Corcos
Catherine Cheze Le Rest
 *INSERM UMR 1101, CHU Morvan
 5 Avenue Foch
 Brest 29200, France
 E-mail: florent.tixier@univ-brest.fr

Published online Oct. 12, 2012.
 DOI: 10.2967/jnumed.112.112573

COPYRIGHT © 2012 by the Society of Nuclear Medicine and Molecular Imaging, Inc.