

## Repeatability of Quantitative Whole-Body $^{18}\text{F}$ -FDG PET/CT Uptake Measures in Patients with Non-Small Cell Lung Cancer: Dynamic Versus Test-Retest Design

**TO THE EDITOR:** In the September issue of *The Journal of Nuclear Medicine*, Kramer et al. comprehensively investigated the repeatability (R) of various quantitative  $^{18}\text{F}$ -FDG uptake metrics in lung cancer patients, including  $\text{SUV}_{\text{max}}$ ,  $\text{SUV}_{\text{mean}}$ ,  $\text{SUV}_{\text{peak}}$ , metabolically active tumor volume (MATV), and total lesion glycolysis (TLG) (1). A test-retest study was performed within 3 d involving double-baseline whole-body  $^{18}\text{F}$ -FDG PET/CT at 60 and 90 min after injection. The results were compared with those of some previously published studies on various oncologic diseases, such as ovarian cancer, non-small cell lung cancer, esophageal cancer, and liver metastases for TLG in particular.

However, a previous study on TLG variability in lung cancer patients—a study involving  $\text{SUV}_{\text{mean}}$  and MATV—was not included in the comparison of Kramer et al. (2). That study assessed R differently; instead of a test-retest acquisition within 3 d over a total of 60 lesions, as used by Kramer et al., a dynamic acquisition involving 10 frames within 60–110 min after injection over 13 lesions was used. Moreover,  $\text{SUV}_{\text{peak}}$  R was obtained from a further dynamic study involving 20 lung cancer lesions (3). We thought it would of interest to compare the two methods—that is, dynamic and test-retest.

The values for dynamic R obtained within 60–110 min after injection versus test-retest R obtained at 90 min after injection for all lesions in the study of Kramer et al. (R is reproducibility coefficient in their Table 3) are as follows: 19.6 versus 23.3%, 14.1 versus 17.8%, 13.2 versus 15.8%, 31.6 versus 23.7%, and 36.4 versus 30.7% for  $\text{SUV}_{\text{max}}$ ,  $\text{SUV}_{\text{mean}}$ ,  $\text{SUV}_{\text{peak}}$ , TLG, and MATV (95% reliability), respectively (1–3). Although the parameter range and the “50% of  $\text{SUV}_{\text{max}}$ ” thresholding method (used for assessing  $\text{SUV}_{\text{mean}}$ , MATV, and hence TLG) were not exactly alike, the dynamic R percentages are consistent with the test-retest R percentages, as much as can be determined considering that R estimate uncertainty was not provided.

Let us further compare the two designs. The dynamic design involved ten 2.5-min frames leading to a  $\pm 12.5$ -min time window around a mean uptake time, in comparison with the test-retest design, which reported a maximal range of 7 min for uptake time (scan 1 in Table 1 (1)). We suggest that reducing the acquisition time for bed positions—that is, less than the 2.5 min that is possible with modern PET/CT systems—and reducing the number of dynamic frames may bridge the two designs. In the framework of assessing response to treatment, it is noteworthy that the dynamic design does not take into account some origins of parameter variability such as changes in plasma glucose level (within the reference range), injected dose, and differences in uptake time. However, Kramer et al. reported that glucose correction does not improve R performance (and even deteriorates it) and that the relative uncertainty about the injected dose is usually very low (1). Regarding differences in uptake time, Kramer et al. showed that the correction proposed by van den Hoff et al. significantly reduced differences between 60-min and 90-min data (4). We would like to emphasize, first, that the time window of the dynamic design includes a variability of  $\pm 12.5$  min around a mean uptake time, which may take into account differences in uptake time usually met in current clinical practice, and second, that simply removing the  $^{18}\text{F}$  physical decay correction can reduce

differences between 60-min and 90-min SUVs (5). Finally, in comparison with the test-retest design, the dynamic design involving several frames reduces the number of lesions to be investigated (and hence the number of patients to be recruited) in order to reliably determine R, since for the same number of lesions, the greater the number of dynamic frames the lower the R estimate uncertainty. We thus suggest that the dynamic design takes into consideration both the patient radiation dose and a busy clinical practice.

In conclusion, the comprehensive study of Kramer et al. about the R value for various  $^{18}\text{F}$ -FDG uptake measures will be useful to nuclear physicians in their current practice (1). That study was achieved using a test-retest design, and we would like to emphasize that the tool box for assessing measurement uncertainty in quantitative PET imaging fortunately offers various designs, each with its own pros and cons. An alternative dynamic design is available that may be particularly suitable when the role of technical parameters in this uncertainty is investigated in an arbitrary PET/CT system. Nevertheless, whatever the design, it clearly appears that guidelines should recommend that any quantitative outcome be accompanied by its measurement uncertainty, which should be specifically determined for each PET/CT system as soon as commissioning.

## REFERENCES

1. Kramer GM, Frings V, Hoetjes N, et al. Repeatability of quantitative whole body  $^{18}\text{F}$ -FDG PET/CT uptake measures as function of uptake interval and lesion selection in non-small cell lung cancer patients. *J Nucl Med*. 2016;57:1343–1349.
2. Laffon E, de Clermont H, Lamare F, Marthan R. Variability of total lesion glycolysis by  $^{18}\text{F}$ -FDG-positive tissue thresholding in lung lesion cancer. *J Nucl Med Technol*. 2013;41:186–191.
3. Laffon E, Burger IA, Lamare F, de Clermont H, Marthan R.  $\text{SUV}_{\text{peak}}$  performance in lung cancer: comparison to average SUV from 40 hottest voxels. *J Nucl Med*. 2016;57:85–88.
4. van den Hoff J, Lougovski A, Schramm G, et al. Correction of scan time dependence of standard uptake values in oncological PET. *EJNMMI Res*. 2014;4:18.
5. Laffon E, de Clermont H, Marthan R. A method of adjusting SUV for injection-acquisition time differences in  $^{18}\text{F}$ -FDG PET imaging. *Eur Radiol*. 2011;21:2417–2424.

Eric Laffon\*

Roger Marthan

\*Hôpital du Haut-Lévêque

avenue de Magellan

33604 Pessac, France

E-mail: elaffon@u-bordeaux2.fr

Published online Dec. 8, 2017.

DOI: 10.2967/jnumed.116.186676

**REPLY:** We thank Drs. Thie, Laffon, and Marthan for their response on our paper (1). In this study, we tried to comprehensively investigate the repeatability of various whole-body  $^{18}\text{F}$ -FDG uptake metrics and assess the influence of several correction methods to normalize  $^{18}\text{F}$ -FDG uptake.

In their letter, Drs. Laffon and Marthan discuss the advantages and disadvantages of a dynamic and test-retest study design for the assessment of  $^{18}\text{F}$ -FDG repeatability. We fully agree that measurement uncertainty of all quantitative PET metrics should be