Journal of Nuclear Medicine, published on February 7, 2019 as doi:10.2967/jnumed.119.226092

Relevance of measurement uncertainty for quantitative response assessment of breast cancer bone metastases with ^{18}F-fluoride.

Eric Laffon [1,2,3 *], Roger Marthan [1,2,3].

[1] CHU de Bordeaux - F-33000 Bordeaux, France.
[2] Univ. Bordeaux, Centre de Recherche Cardio-Thoracique de Bordeaux, F-33000 Bordeaux, France.
[3] INSERM U-1045, Centre de Recherche Cardio-Thoracique de Bordeaux F-33000 Bordeaux, France.

*Correspondence to Dr Eric Laffon, Service de Médecine Nucléaire, Hôpital du Haut-Lévèque, avenue de Magellan, 33604 PESSAC, France.
Telephone: +33 5 57 65 68 38; Fax: +33 5 57 65 68 39;
E-mail: elaffon@u-bordeaux.fr

TO THE EDITOR: To assess treatment response of bone metastases in breast cancer by positron emission tomography/ computed tomography (PET/CT), Dr Azad et al. recently compared the change in Ki (^{18}F-fluoride metabolic flux to bone mineral) with that in maximum and mean standardized uptake values (SUV_{max}, SUV_{mean} assessed with a 40% of the SUV_{max} thresholding method) *(1)*. Calculation of Ki in individual bone metastases was based on a semi-population input function (IF), which can be considered as a further development of the Hunter's approach published for ^{18}F-FDG, and whose equation of its "residual" part for ^{18}F-fluoride was, unfortunately, not provided *(2)*. On the ^{18}F-fluoride PET/CT, the authors defined progressive disease (PD) as an ≥ 25% increase in Ki, SUV_{max} or SUV_{mean}, whereas non-PD included partial responders (> 25% decrease in Ki, SUV_{max} or SUV_{mean}) and stable disease (<25% increase or decrease). This classification was adapted from European Organization for Research and Treatment of Cancer criteria, acknowledging it was originally described for ^{18}F-FDG *(3)*. The conclusion was that, after 8 weeks of endocrine treatment for bone-predominant metastatic breast cancer, Ki more reliably differentiated PD from non-PD than SUV_{max} and SUV_{mean}, probably

because measurement of SUVs underestimates fluoride clearance as changes in input function are not accounted for.

We believe that using the same ≥ 25% increase to define PD (and, consequently, non-PD) for Ki as well as both SUV_{max} and SUV_{mean} should be questioned, since each of these quantitative parameters has its own measurement uncertainty (MU) *(4)*. This issue is not related to the fact that the 25% threshold was originally described for ^{18}F-FDG and not ^{18}F-fluoride since MU of an arbitrary metrics is very likely close for several ^{18}F-labelled tracers *(5)*. In fact, MU of a quantitative parameter includes its repeatability (R), defined as the minimal relative change between two measurements assessed from two scans carried out under changed conditions of measurement that is required to consider a significant difference (called "reproducibility" by the Joint Committee for Guides in Metrology *(4)*). This definition clearly highlights its relevance for assessing treatment response in solid tumors. The repeatability magnitude for SUV_{max} and SUV_{mean} assessed with a 40% of the SUV_{max} thresholding method, may be estimated to be 19.6 and 13.8% (95% confidence level), respectively, from previously published studies performed in lung cancer patients with ^{18}F-FDG, illustrating that averaging leads to lowering MU *(6,7)*. In contrast, to the very best of our knowledge, repeatability of Ki expressed in Equation 3 of supplementary data by Azad et al. *(1)* remains unknown. However, one can ascertain that it is much greater than that of SUV_{max} or SUV_{mean}, and even greater than 25%, because, besides sharing the MU of the tracer concentration in the bone region of interest with SUV_{mean}, it additionally combines MU of the tracer concentration in the plasma, MU of the distribution volume in the unbound bone pool (arbitrary set by the authors as that of a population mean value), and MU of the normalized time "Θ(T)" that requires adjusting the tracer release rate constant "k_{loss}" (Equation 3 by Azad et al.). In other words, whereas an ≥ 25% increase in SUV_{max} for defining PD and non-PD is consistent with SUV_{max} repeatability (about 19.6%), the same 25% value for Ki (whose repeatability is very likely much greater than 25%) and for SUV_{mean} (whose repeatability is about 13.8%) may not be appropriate. Finally, we would like to emphasize that the use of an average SUV, or a derived metrics, that can be obtained by pooling several hottest voxels possibly located in several separate places over the whole skeleton, by reducing repeatability, might thus be a relevant quantitative tool at the patient level for predicting response to treatment *(8)*.

To conclude, Dr Azad et al. relevantly put forward the interest of assessing treatment

response of bone metastases in breast cancer by ^{18}F-fluoride PET imaging. We further suggest that taking into account the repeatability of different metrics to define PD (and non-PD), rather than using the same percentage value of relative increase (or decrease), may prove of clinical value.

REFERENCES

1. Azad G, Siddique MM, Taylor B et al. Does measurement of ^{18}F-fluoride metabolic flux improve response assessment of breast cancer bone metastases compared with standardized uptake values in ^{18}F-fluoride PET/CT? *J Nucl Med.* 2018 Jul 24. pii: jnumed.118.208710. doi: 10.2967/jnumed.118.208710. [Epub ahead of print].

2. Hunter GJ, Hamberg LM, Alpert NM, Choi NC, Fischman AJ. Simplified measurement of deoxyglucose utilization rate. *J Nucl Med*. 1996;37:950–955.

3. Lecouvet FE, Talbot JN, Messiou C, Bourguet P, Liu Y, de Souza NM. Monitoring the response of bone metastases to treatment with Magnetic Resonance Imaging and nuclear medicine techniques: a review and position statement by the European Organisation for Research and Treatment of Cancer imaging group. *Eur J Cancer*. 2014;50:2519–2531.

4. JCGM 2008. Evaluation of measurement data - Guide to the expression of uncertainty in measurement. www.bipm.org, September 2008.

5. Laffon E, de Clermont H, Buj S, Marthan R. Should we assess repeatability of PET quantitative uptake measurements of each ^{18}F-labelled tracer? *Eur J Nucl Med Mol Imaging.* 2018;45:1272–1273.

6. Laffon E, de Clermont H, Lamare F, Marthan R. Variability of total lesion glycolysis by ^{18}FDG-positive tissue thresholding in lung cancer. *J Nucl Med Technol*. 2013.41:186–191.

7. Laffon E, Lamare F, de Clermont H, Burger IA, Marthan R. Variability of average SUV from several hottest voxels is lower than that of SUVmax and SUVpeak. *Eur Radiol.* 2014;24:1964–1970.

8. Laffon E, de Clermont H, Marthan R, Paycha F. On the ^{18}F-fluoride PET imaging quantification to predict 223Ra-dichloride treatment response. *Eur J Nucl Med Mol Imaging.* 2018;45:318–320.