abiraterone or enzalutamide. Further, all patients in VISION had progressed after docetaxel and approximately 40% had progressed after cabazitaxel. Thus, the patient populations of ALSYMPCA and VISION are completely distinct. Indirect comparisons between phase III trials are always fraught with difficulty. In this case, because the populations are so distinct, comparisons would be particularly problematic.

Dr. Duarte also raises the issue that the liver metastasis patients do not have improved survival in VISION and suggests that the positive effects of ^{177}Lu-PSMA-617 may be predominantly on patients with bone metastases. Although these points are well taken, the overanalysis of small data subsets can at times be erroneous. The number of patients in the VISION trial with liver metastases was far smaller than optimal for a conclusive analysis. There is much more to learn before a definitive conclusion can be drawn. Further, we would all agree that there is considerable heterogeneity for those with liver metastases and that more analyses may potentially yield interesting findings. Perhaps the patients with higher PSMA PET SUVs may be distinct from those with lower PSMA PET SUVs. Perhaps those with more than 20 liver metastases may be distinct from those having just one. Simply stated, there is much more to learn before categoric statements can be made regarding analyses of underpowered subsets.

REFERENCES

1. Sartor O, de Bono J, Chi KN, et al. Lutetium-177-PSMA-617 for metastatic castration resistant prostate cancer. *N Engl J Med.* 2021;385:1091–1103.
2. Parker C, Nilsson S, Heinrich D, et al. Alpha emitter radium-223 and survival in metastatic prostate cancer. *N Engl J Med.* 2013;369:213–223.

Oliver Sartor
Tulane University School of Medicine
E-mail: osartor@tulane.edu

DOI: 10.2967/jnumed.121.263193

On Semiquantitative Methods for Assessing Vascular ^{18}F-FDG PET Activity in Large-Vessel Vasculitis

TO THE EDITOR: In a series of 95 large-vessel vasculitis patients investigated with ^{18}F-FDG PET imaging, Dashora et al. recently tested the performance of qualitative (PET vascular activity score [PETVAS]) and semiquantitative (SUV and tissue-to-background ratio [TBR] relative to liver and blood activity) scoring methods (*1*). Regarding the latter methods, 9 territories were created in each patient by segmenting the aorta and branch arteries. A territory score was calculated by averaging the SUV_{max} assessed in each axial region of interest that was manually drawn across the territory, and a global summary, SUV_{Artery}, was then calculated by averaging all territory scores. Liver TBR (TBR_{Liver}) and blood TBR (TBR_{Blood}) were computed by dividing SUV_{Artery} by a mean liver and blood SUV, respectively. The performance of each metric was assessed in association with reader interpretation of vascular PET activity and with physician assessment of clinical disease activity, including the area under the receiver-operating-characteristic curve. Tables 2 and 3 by Dashora reported the metrics performance against the 2 reference standards; this performance was poor–poor for SUV_{Artery} (area under receiver-operating-characteristic curve, 0.67–0.59) and good–poor for TBR_{Liver} and PETVAS (areas under receiver-operating-characteristic curve, 0.85–0.66 and 0.87–0.65, respectively) (*1*). TBR_{Blood} had slightly lower performance than TBR_{Liver}.

Since TBR_{Liver} involves SUV_{Artery}, which results from SUV_{max} averaging, we suggest that instead of using SUV_{Artery}, we use an averaged SUV_{max} obtained from N hottest voxels ($SUV_{max\text{-}N}$) irrespective of their location within the 9 vascular territories (*2*). Both SUV_{Artery} and $SUV_{max\text{-}N}$ take into consideration the heterogeneity of the vessel-wall uptake, but N can actually be much greater than the total number of regions of interest used by Dashora et al. for calculating SUV_{Artery}. Since the greater the N number, the lower the $SUV_{max\text{-}N}$ variability, a more reliable TBR_{Liver} can thus be provided than with SUV_{Artery} (*2*,*3*). A previous assessment of treatment response in a Takayasu arteritis patient illustrates the possible magnitude of N, with $SUV_{max\text{-}N}$ pooling $N = 4{,}100$ and 515 voxels, corresponding to a hottest volume V = 100 and 12.6 mL, respectively (*4*). $SUV_{max\text{-}V}$ might be preferred to $SUV_{max\text{-}N}$, for the voxel volume depends on the PET system at a given center. For assessing response to treatment in a large-vessel vasculitis patient, it has been previously shown that V (or N) should be set in the scan showing the lowest total ^{18}F-FDG–positive volume, which is expected to be posttreatment one (*4*). For assessing the severity of large-vessel vasculitis inflammation as in the study of Dashora et al., we suggest that standard $SUV_{max\text{-}V}$–based TBR_{Liver} metrics might be relevant, using an arbitrary value of V defined by expert consensus (e.g., of 10 cm^3). Additionally, we suggest that the hottest volume V corresponding to a standard value of $SUV_{max\text{-}V}$–based TBR_{Liver} could also be investigated by Dashora et al. as a further metric. This TBR_{Liver} standard value should be greater than 1, as is consistent with the qualitative territory score of 3 used in PETVAS (arterial uptake > liver uptake). The standard might be set at 1.33 according to TBR_{Liver} data reported in Table 3 by Dashora et al. for physician assessment of clinical disease activity, that is, between the clinical-active range and the clinical-remission range ($1.33 = 1.27 + 1.96 \times 0.03 \approx 1.46–1.96 \times 0.06$) (*1*). A similar line of argument provides a TBR_{Blood} standard value of 2.43 (from Table 3 of Dashora et al. (*1*)).

To conclude, we fully agree with the authors that qualitative metrics for assessing large-vessel vasculitis inflammation severity with ^{18}F-FDG PET, such as PETVAS, are attractive in clinical practice because of ease of implementation and ease of interpretation. However, we believe that $SUV_{max\text{-}V}$–based TBR_{Liver} (or $SUV_{max\text{-}V}$–based TBR_{Blood}) could also be used daily if manufacturers are encouraged to make $SUV_{max\text{-}V}$ (or $SUV_{max\text{-}N}$) easier to assess than currently (*2*–*4*).

REFERENCES

1. Dashora HR, Rosenblum JS, Quinn KA, et al. Comparing semi-quantitative and qualitative methods of vascular FDG-PET activity measurement in large-vessel vasculitis. *J Nucl Med.* June 4, 2021 [Epub ahead of print].
2. 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.
3. Laffon E, Burger IA, Lamare F, de Clermont H, Marthan R. SUVpeak performance in lung cancer: comparison to average SUV from the 40 hottest voxels. *J Nucl Med.* 2016; 57:85–88.
4. Caubet O, Meunier V, Marthan R, Laffon E. Early assessment of treatment response in Takayasu arteritis: an ^{18}FDG PET procedure. *Clin Nucl Med.* 2016;41:743–745.

Eric Laffon*
Roger Marthan
**CHU de Bordeaux, University of Bordeaux, and INSERM U-1045, Bordeaux, France*
E-mail: elaffon@u-bordeaux.fr

Published online Sep. 2, 2021.
DOI: 10.2967/jnumed.121.263060

Reply: On Semiquantitative Methods for Assessing Vascular ^{18}F-FDG PET Activity in Large-Vessel Vasculitis

REPLY: We were pleased that discussion was brought forth by Laffon and Marthan because of our recent paper on quantitative and qualitative ^{18}F-FDG PET for large-vessel vasculitis (LVV) (*1*). Indeed, we agree it could be revealing to attempt a measurement strategy that involves the appreciation of the hottest *N* number of voxels ($SUV_{max\text{-}N}$) as proposed by the authors. If *N* is greater than the number of single SUV_{max} measurements from each region of interest drawn over the entire arterial tree, $SUV_{max\text{-}N}$ could lead to an overall more reproducible value in addition to a potentially greater contribution of abnormal activity in regions of active vasculitis than in regions without inflammation.

Although specific methodology to quantify vascular inflammation will no doubt be tested and refined, we would like to emphasize our underlying thought process for the design of our quantitative methodology, with reference to how this and other strategies for quantitative PET might be used for LVV. We will organize our discussion around 3 questions: What can be deployed clinically? What is most useful in clinical trials? What are we trying to do with vascular imaging in LVV?

For clinical deployment (question 1), even with the recent advent of greater acceptance of ^{18}F-FDG PET in clinical evaluation of inflammatory disease (*2*), we acknowledge that large-vessel vasculitis is a rare disease that many interpreting physicians will not encounter frequently. Our experience is that extensive familiarity and care are necessary to rigorously apply a complex quantitative strategy that involves contouring of the arteries as applied in this study, which did not have the advantage of intravenous contrast medium for guidance. Regardless of how the specific voxels are aggregated mathematically, the contouring itself is likely to be beyond the abilities of the standard medical professional in routine clinical practice. Hence, our introduction of a qualitative metric such as PETVAS (*3*), which is similar, but not identical, to the emerging use of an ordinal scoring system in lymphoma (*4*). We showed that PETVAS is a reasonable clinically deployable alternative to what we felt was an inevitable question from the community, which is "why not use SUVs?" Another compelling reason to not yet favor the use of metrics such as SUV in the clinic for LVV relates to the common misapplication of quantitative metrics from the literature for sensitivity and specificity in image interpretation. The performance characteristics of a quantitative metric are appropriately applied if images can be reproduced in a uniform format, which must be standardized across vendors with identical imaging characteristics that harmonize important features such as resolution, noise, voxel size, and postreconstruction filtering. Despite recent meaningful attempts (*5*), such a level of uniform standardization will likely not soon be achievable in clinical practice.

For clinical trials (question 2), we see a role for complementary advanced quantitative strategies as we and others have proposed. Clinical trials more often involve multiple imaging time points of the same subject before and after a treatment or intervention, using the same imaging characteristics. Our project highlighted that both qualitative and quantitative methods are associated with clinical measures of disease activity, and both approaches could be used to facilitate discovery in research; however, qualitative approaches potentially offer more precision and reliability.

Regarding question 3, it may sound odd to ask "what are we actually trying to do?" As investigators conducting an ongoing, large prospective observational cohort study on LVV, we would like to emphasize that interpretation of ^{18}F-FDG PET findings should be considered in the context of disease activity assessment across other domains. ^{18}F-FDG PET is only 1 facet of the multidisciplinary approach needed to fully realize patient-specific treatment guidance. Comprehensive clinical, laboratory, and imaging assessment is often helpful to accurately assess disease activity and inform management decisions. The cumulative burden of vascular involvement does not always correlate with clinical outcomes. A small focal inflammatory lesion in a single artery may lead to severe vascular damage with disastrous consequences, whereas profound near pan-arterial intense inflammation may occur in an otherwise asymptomatic patient. To inform the details of a better qualitative or quantitative evaluation for individualized care with advanced methods proposed by our group or others, we must continue to define the complex associations between ^{18}F-FDG PET findings and clinical outcomes in LVV. Controlled environments, such as randomized clinical trials, will go further to answer questions related to the combinatory use of qualitative and quantitative PET, as well as specifics for the production of each.

DISCLOSURE

This study was supported by the Intramural Research Programs of the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS). No other potential conflict of interest relevant to this article was reported.

REFERENCES

1. Dashora HR, Rosenblum JS, Quinn KA, et al. Comparing semi-quantitative and qualitative methods of vascular FDG-PET activity measurement in large-vessel vasculitis. *J Nucl Med.* June 4, 2021 [Epub ahead of print].
2. Wahl RL, Dilsizian V, Palestro CJ. At last, ^{18}F-FDG for inflammation and infection! *J Nucl Med.* 2021;62:1048–1049.
3. Grayson PC, Alehashemi S, Bagheri AA, et al. ^{18}F-fluorodeoxyglucose–positron emission tomography as an imaging biomarker in a prospective, longitudinal cohort of patients with large vessel vasculitis. *Arthritis Rheumatol.* 2018;70:439–449.
4. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol.* 2014;32:3059–3068.
5. Slart RHJA, writing group; reviewer group; members of EANM Cardiovascular; members of EANM Infection & Inflammation; Members of Committees, SNMMI Cardiovascular; Members of Council, PET Interest Group; Members of ASNC; EANM Committee Coordinator. FDG-PET/CT(A) imaging in large vessel vasculitis and polymyalgia rheumatica: joint procedural recommendation of the EANM, SNMMI, and the PET Interest Group (PIG), and endorsed by the ASNC. *Eur J Nucl Med Mol Imaging.* 2018;45:1250–1269.

Mark A. Ahlman
Roberto Maass-Moreno
Peter C. Grayson*
**National Institutes of Health*
Bethesda, Maryland
E-mail: peter.grayson@nih.gov

Published online Oct. 28, 2021.
DOI: 10.2967/jnumed.121.263158