

**Comparing Semi-quantitative and Qualitative Methods of Vascular FDG-PET Activity
Measurement in Large-Vessel Vasculitis**

Himanshu R. Dashora BS ¹, Joel S. Rosenblum BS ¹, Kaitlin A. Quinn MD ¹, Hugh Alessi BS ¹,
Elaine Novakovich RN ¹, Babak Saboury MD ², Mark A. Ahlman MD ^{2*}, Peter C. Grayson
MD, MSc ^{1*}

¹Systemic Autoimmunity Branch, NIAMS, National Institutes of Health

²Radiology and Imaging Sciences, Clinical Center, National Institutes of Health

**Equal contribution*

Running Title: FDG-PET in Vasculitis

Financial Support: This work was supported through the intramural research program at the
National Institute of Arthritis Musculoskeletal and Skin Disease (ZIA-AR-041199)

Word Count: 4,736

Corresponding Author:

Peter C. Grayson
National Institutes of Health
9000 Rockville Pike, Bethesda, MD 20892
Email: peter.grayson@nih.gov
Phone: 301-827-9187
Fax: 301-480-3978

ABSTRACT

The study rationale was to assess the performance of qualitative and semi-quantitative scoring methods for ^{18}F -fluorodeoxyglucose positron emission tomography (FDG-PET) assessment in large-vessel vasculitis (LVV).

Methods

Patients with giant cell arteritis (GCA) or Takayasu's arteritis (TAK) underwent clinical and imaging assessment, blinded to each other, within a prospective observational cohort. FDG-PET-CT scans were interpreted for active vasculitis by central reader assessment. Arterial FDG uptake was scored by qualitative visual assessment using the PET vascular activity score (PETVAS) and by semi-quantitative assessment using standardized uptake values (SUV) and target-to-background ratios (TBR) relative to liver/blood activity. Performance of each scoring method was assessed by intra-rater reliability using the intra-class coefficient (ICC) and area under receiver-operator characteristic curves (AUC), using physician assessment of clinical disease activity and reader interpretation of vascular PET activity as independent reference standards. Wilcoxon signed-rank test was used to analyze change in arterial FDG uptake over time.

Results

Ninety-five patients (GCA=52; TAK=43) contributed 212 FDG-PET studies. The ICC for semi-quantitative evaluation [0.99 (range 0.98-1.00)] was greater than the ICC for qualitative evaluation [0.82 (range 0.56-0.93)]. PETVAS and TBR metrics were more strongly associated with reader interpretation of PET activity than SUV metrics. All assessment methods were

significantly associated with physician assessment of clinical disease activity, but the semi-quantitative metric TBR_{Liver} achieved the highest AUC (0.66). Significant but weak correlations with C-reactive protein were observed for SUV metrics ($r = 0.19$, $p < 0.01$) and TBR_{Liver} ($r = 0.20$, $p < 0.01$) but not for PETVAS. In response to increased treatment in 56 patients, arterial FDG uptake was significantly reduced when measured by semi-quantitative (TBR_{Liver} 1.31 to 1.23, 6.1% Δ , $p < 0.0001$) or qualitative (PETVAS 22 to 18, $p < 0.0001$) methods. Semi-quantitative metrics provided complementary information to qualitative evaluation in cases of severe vascular inflammation.

Conclusion

Both qualitative and semi-quantitative methods to measure arterial FDG uptake are useful to assess and monitor vascular inflammation in LVV. Compared to qualitative metrics, semi-quantitative methods have superior reliability and better discriminate treatment response in cases of severe inflammation.

Key Words

Vasculitis, Takayasu's arteritis, giant cell arteritis, PET, outcome measure

INTRODUCTION

Large-vessel vasculitis (LVV) refers to a class of rare diseases characterized by inflammation of the aorta and its primary branch arteries. Giant cell arteritis (GCA) and Takayasu's arteritis (TAK) comprise the two major subtypes of LVV (1). ^{18}F -fluorodeoxyglucose (FDG) positron emission tomography (PET) can detect metabolic activity in the walls of large arteries as a biomarker of vascular inflammation (2). Ample evidence supports the use of FDG-PET as a diagnostic surrogate to histologic confirmation of vasculitis, which is advantageous because arterial biopsies are invasive and can be difficult to obtain (3,4). In contrast to diagnostic assessment, use of arterial FDG uptake to guide treatment decisions and monitor disease activity is less well defined (5-9), in part due a lack of prospective, longitudinal imaging studies in LVV (10,11). Reliance on clinical assessment alone may lead to under detection of vascular pathology (12). Vascular inflammation with angiographic progression of disease can occur in patients with LVV who are otherwise completely asymptomatic, highlighting a need for vascular imaging to complement clinical assessment in these patients (13).

Uncertainty about the optimal method to evaluate FDG uptake in the large arteries remains a major barrier to the use of FDG-PET to monitor vascular inflammation (11). Both visual/qualitative and semi-quantitative methods of FDG-PET assessment have been reported in LVV. Qualitative methods typically visually compare the amount of FDG uptake in the arterial wall relative to a background tissue, such as the liver (11,14), similar to the Deauville score used in lymphoma (15). In contrast, semi-quantitative methods use regions of interest (ROI) constructed on the PET image to determine maximum standard uptake values (SUV) (16). Target-to-Background Ratios (TBR), comprised of SUVs from arterial tissue referenced to background tissue (e.g. liver, blood pool), are also used to quantify arterial FDG uptake in

atherosclerosis and vasculitis (17). Recent recommendations highlight that several methods to quantify arterial FDG uptake are available but the relevance of each method to evaluate patients requires further clarification (11). SUV metrics often overlap between patients with LVV and controls, and many patients with LVV have residual, and sometimes profound, arterial FDG uptake during periods of apparent clinical remission (18).

There is an unmet need to better understand the strengths and weakness of qualitative versus semi-quantitative methods to quantify arterial FDG uptake in LVV. Semi-quantitative assessment of arterial FDG uptake evaluation can be a time-consuming process, which may be difficult to apply in a contemporary clinical setting or be cost-prohibitive in research. In contrast, qualitative PET assessment may be easier to do with appropriate user training; however, qualitative assessment may be less reliable and accurate to quantify arterial FDG uptake compared to semi-quantitative approaches (19,20).

This study aimed to compare the effectiveness of qualitative and semi-quantitative scoring methods with the goal of informing a standardized approach to FDG-PET assessment in LVV for use in clinical care and research.

MATERIALS AND METHODS

Study Population

Patients, age ≥ 18 years, with LVV were recruited into a prospective, observational cohort at the National Institutes of Health in Bethesda, MD. All patients provided written informed consent, and the study was approved by an institutional review board at the National Institutes of Health (NCT02257866; 14-AR-0200). All patients fulfilled the 1990 American College of Rheumatology (ACR) Classification Criteria for TAK (21) or modified 1990 ACR Criteria for GCA (22,23). Patients were enrolled at various stages of the disease course. Treatment decisions were made at the discretion of each patient's local health care provider rather than the investigative research team.

Clinical Assessment

All patients underwent clinical and imaging assessments within a 24-hour time period at the National Institutes of Health Clinical Center. Repeat imaging studies and clinical assessments were performed at 6-month intervals. A team of clinical rheumatologists with further specialist training and experience in LVV evaluated all cases. Physician assessment of clinical disease activity was recorded as *active* or *remission* based on findings from the medical history, physical examination, and laboratory assessments. Active disease was defined as the presence of clinical disease features attributed to vasculitis (e.g. carotidynia) at the time of assessment. Remission was defined as the absence of clinical symptoms attributable to vasculitis at the time of assessment. Imaging study findings were not incorporated into the definition of clinical disease activity.

FDG-PET Imaging Protocol

All patients underwent FDG-PET computed tomography (CT) on a 128 detector-row Siemens Biograph mCT (Siemens Medical Solutions, Erlangen, Germany). Patients were given detailed instructions to avoid carbohydrate-laden meals one day prior to imaging and to fast on the day of imaging. ^{18}F -FDG dosage was fixed to 370 MBq for all patients. Image acquisition of the torso was performed two hours after injection. Post-acquisition image reconstruction utilized CT attenuation correction and iterative reconstruction (point spread function correction with time of flight, 3 iterations, 21 subsets, 256 matrix, final isotropic voxel resolution of 3.2mm^3 , no post-reconstruction filtering).

FDG-PET Imaging Assessment

Qualitative Analysis. One imaging specialist (MA) interpreted all PET studies, blinded to clinical data. A study was excluded if there were technical concerns about image quality per physician review. Each study was subjectively interpreted as *PET Active* or *PET Inactive* if there was at least one area of abnormal arterial FDG uptake felt to represent vascular inflammation. Intra- and inter-rater reproducibility of LVV PET image interpretation of the specialist has been previously reported to be excellent (18). Qualitative assessment of FDG uptake was also performed at the territory level; which included 4 segments of the aorta (ascending, arch, descending thoracic, and abdominal) and 5 branch arteries (brachiocephalic, right and left carotid, right and left subclavian). Scores between 0 and 3 were assigned to each territory, representing the visual degree of arterial FDG uptake relative to liver FDG uptake (0 = no uptake, 1 = less than liver, 2 = similar to liver, 3 = greater than liver). Adding the qualitative arterial territory scores yields a summary score (termed the PET Vascular Activity Score,

PETVAS) ranging from 0-27; with higher scores indicating a greater global burden of vascular inflammation (18).

Semi-Quantitative Analysis. ROIs were manually contoured in OsiriX DICOM Viewer (Version 9.5.2, Bernex, Switzerland) with respect to both CT anatomic location and co-registered PET activity to determine arterial FDG uptake values. ROIs were drawn in the axial dimension, encompassing both arterial wall and lumen. Five segments of the aorta (ascending aorta, aortic arch, descending thoracic aorta, suprarenal abdominal aorta, infrarenal abdominal aorta), and four branch arteries (right and left common carotid and subclavian arteries) were segmented in this process to create nine territories. The maximum FDG uptake values per ROI (SUV_{max}) of each territory were identified. A territory score was calculated by taking the average of the SUV_{max} across all ROIs in the territory (17). A global summary metric (SUV_{Artery}) was calculated by averaging all territory scores.

Volumetric mean FDG uptake activity in the liver (SUV_{Liver}) was measured in the dome of the right lobe. Volumetric mean venous blood pool activity (SUV_{Blood}) was measured within the right jugular, superior vena cava, right atrium, and the inferior vena cava. SUV_{Artery} was divided by the background tissue to generate two tissue-to-background ratio (TBR) metrics:

TBR_{Liver} , TBR_{Blood} .

Statistical Analysis

Intra-rater Reliability. Intra-rater reliability (IRR), reflecting the variation in data measured by one rater over multiple trials, was quantified with a two-way random effect, consistency, single measurement intra-class coefficient (ICC) (24). ICC estimates and their 95% confident

intervals were calculated using R-package irr (Version 0.84.1, Matthias Gamer). ICC values lie between 0 and 1. Values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9 are indicative of poor, moderate, good, and excellent reliability, respectively. ICC for the qualitative approach was obtained by repeating PETVAS on a set of randomly selected patients. ICC for the semi-quantitative approach was obtained by recontouring ROIs to recalculate SUV_{Artery} for a set of randomly selected patients representing 10% of the cohort.

Receiver-Operating Characteristic (ROC) Curve. Area under the ROC curve (AUC) along with the 95% confidence interval (CI) was utilized as a combined measure of sensitivity and specificity to evaluate the overall performance of the PET scoring metrics as classifiers of a binary outcome (25), either reader interpretation of vascular PET activity (*PET Active* vs. *PET Inactive*) or physician assessment of clinical disease activity (*Clinical Active* vs. *Clinical Remission*). AUC values lie between 0 and 1. Metrics with capability to distinguish between binary outcomes will result in an AUC above 0.5, with larger AUC values suggesting better diagnostic performance. The Youden's J statistic was used to determine the optimal cut-off score that maximized the distance to the identity (diagonal) line.

Mixed Effects Logistic Regression. To account for repeated imaging contributions from a single patient, generalized linear mixed models (GLMMs) with logistic outcomes were constructed. The dependent variable was a binary classification of either reader interpretation of vascular PET activity (*PET Active* vs. *PET Inactive*) or physician assessment of clinical disease activity (*Clinical Active* vs. *Clinical Remission*). The PET scoring metric, either semi-quantitative or qualitative, was used as the fixed effect with patient identification as a random

effect. A Bound Optimization by Quadratic Approximation (BOBYQA) nonlinear optimizer and 10 points of integration for the adaptive Gaussian Hermite approximation were used as model control parameters. Independent GLMMs were created for each scoring method. The Akaike information criterion (AIC) estimates the information loss for a given model and is a means for model selection. Relative to the other models, the candidate model with the lowest AIC minimizes estimated information loss. All GLMM analysis was performed using R-package lme4 (Version 1.1-21).

Correlation Analysis. Spearman's rank order correlation was used to measure the association between the PET scoring metrics and acute phase reactants: C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR). The Spearman's r , ranging from 0 to 1, and p value of correlation are presented.

Longitudinal PET Assessment in Response to Treatment. The Wilcoxon matched pairs signed rank test was used to compare change in PET assessment metrics between two time points for the same patient. When stratifying by treatment status, initial and follow-up scan pairings were placed into *increased treatment* and *no change* groups. Increased treatment was defined by the introduction of a glucocorticoid sparing medication or an increase in daily prednisone dose by > 5mg. No change was defined by the maintenance of biologic agent administration or stable glucocorticoid dosage.

Semi-quantitative metrics of FDG-PET activity exist on a continuous scale. In contrast, a qualitative metric like PETVAS is ordinal with a maximum score of 27 (18). In cases of severe inflammation where PET activity may be reduced but remain in a range above the maximum

PETVAS score, semi-quantitative metrics may be better suited to demonstrate change in PET activity. A subset of patients was selected who had 1) severe vascular inflammation defined by a baseline PETVAS score of 27 and 2) reduction in FDG uptake on the follow-up scan by visual assessment with agreement by two independent readers. Longitudinal change in PET activity measured by PETVAS versus TBR_{Liver} metrics was compared in this subset of patients.

RESULTS

Study Population

A total of 95 patients (GCA=52; TAK=43) contributed 212 imaging studies. Three imaging studies were excluded due to concerns about image quality. Demographics were consistent with the expected age and sex distributions for GCA and TAK (Table 1). Patients were seen on average 6.1 years into the disease course while taking on average 8.3 mg of daily prednisone.

Intra-rater Reliability

Intra-rater reliability for repeat scoring of 34 imaging studies using the semi-quantitative (SQ) scoring protocol was excellent (ICC = 0.99, range 0.98 – 1.00). Intra-rater reliability for the qualitative assessment by PETVAS was good (ICC = 0.82, range 0.56 – 0.93).

Quantification of Arterial FDG Uptake in Association with Reader Interpretation of PET Scan Activity

Out of 209 FDG-PET imaging studies, 147 scans were interpreted as *PET Active* and 62 scans as *PET Inactive*. Compared to the use of SUV alone, discriminatory power (AUC) was greater and model quality was better (lower AIC) when TBR was used to differentiate *PET Active* from *PET Inactive* scans (Table 2). PETVAS performed similarly well compared to TBR, with better performance characteristics than SUV. PETVAS achieved the highest AUC, and lowest AIC relative to the other models, with an optimal cutoff of 19.5 (Table 2).

Quantification of Arterial FDG Uptake in Association with Physician Assessment of Clinical Disease Activity and Laboratory Tests

Complete clinical and imaging assessments were available for 206 study visits. Clinical disease activity was assessed as *Clinical Active* for 95 study visits and *Clinical Remission* for 131 study visits. Corresponding arterial FDG uptake evaluated by any proposed method significantly discriminated active disease from clinical remission, but TBR metrics and PETVAS resulted in higher AUC values than SUV metrics (Table 3). Within the proposed mixed models, the PETVAS-informed model had the lowest AIC when predicting the same clinical outcomes as the other models, suggesting the best model fit (26). Broadly, all AUC values were lower when FDG metrics were compared with clinical assessment compared to reader interpretation of PET activity as the reference standard.

Significant but weak correlations with acute phase reactants were observed for SUV_{Artery} (CRP $r = 0.19$, $p < 0.01$; ESR $r = 0.14$, $p = 0.04$) and TBR_{Liver} (CRP $r = 0.20$, $p < 0.01$; ESR $r = 0.15$, $p = 0.03$). Neither TBR_{Blood} or PETVAS was significantly correlated with CRP or ESR (Table 4).

Longitudinal Treatment Response

Treatment was increased over 56 interval study visits. Corresponding, there was significant reduction in vascular inflammation by semi-quantitative (TBR_{Liver} Median (IQR): 1.31 (1.19-1.59) to 1.23 (1.13-1.39), $p < 0.001$) or qualitative approaches (PETVAS Median (IQR): 22 (17-25) to 18 (15-22), $p < 0.001$). Over 25 interval visits for which there was no change in treatment status between successive imaging studies, the degree of vascular inflammation remained similarly unchanged as measured by either semi-quantitative (TBR_{Liver} Median (IQR): 1.39 (1.24-1.54) to 1.35 (1.78-1.49), $p = 0.22$ or qualitative (PETVAS Median (IQR): 21 (18-25) to 21 (18.5-25), $p = 0.68$) assessments (Figures 1, 2).

A subset of nine patients with severe inflammation (baseline PETVAS of 27) who had a visually apparent reduction of arterial FDG-uptake on the follow up imaging study were studied. PETVAS was significantly reduced from a baseline score of 27 to a median score of 24 (18.5 – 26), $p < 0.01$ at the follow up visit (Figure 3). TBR_{Liver} scores in these same patients were a median of 1.86 (1.55 - 2.63) at the baseline visit with a significant reduction of scores at follow up (median = 1.24, range = 1.14 – 1.69, $p < 0.01$). While the baseline PETVAS scores were the same for all 9 patients, there was a corresponding dynamic range of baseline TBR_{Liver} scores, reflecting variability among these patients. TBR_{Liver} was reduced over time in every patient; however, there was minimal reduction of PETVAS (i.e. change ≤ 1 point) in 3 of 9 patients. Representative images from a patient with visually apparent reduction in vascular PET activity are shown in Figure 4.

DISCUSSION

Use of FDG-PET to monitor vascular inflammation in LVV holds promise as a complement to clinical and laboratory-based assessment (10,18,27,28). Visualizing glucose metabolism within the arterial wall as a biomarker of vascular inflammation enables clinicians to non-invasively diagnose and track disease activity in LVV directly in the target tissue, in parallel with clinical and laboratory assessments (29). This is particularly important in LVV because patients can develop subclinical vascular inflammation that can only be detected and monitored by vascular imaging without accompanying clinical symptoms or abnormal laboratory findings (18,28,30). The present study advances our understanding about the strengths and weaknesses of different methodologic approaches to quantify vascular inflammation.

Reassuringly, both qualitative and semi-quantitative approaches performed well to detect and monitor arterial FDG PET uptake in patients with LVV. PETVAS, a qualitative scoring approach developed by our group, and semi-quantitative methods had good to excellent intra-rater reliability. Because some patients can have vascular inflammation by PET in absence of clinical activity, we studied performance characteristics of qualitative and semi-quantitative metrics against two independent reference standards (31). As expected, SUV metrics, TBRs, and PETVAS were each significantly associated with reader interpretation of vascular PET activity; however, TBRs and PETVAS outperformed SUV metrics as evidenced by higher AUC in the models. When compared against physician assessment of clinical disease activity as the reference standard, all the metrics distinguished between active clinical disease and remission with lower AUC values than when using reader interpretation of PET activity as the reference standard, reflecting that clinical assessment is not always linked to vascular inflammation. Both

qualitative and semi-quantitative approaches were useful to demonstrate reduction in the burden of vascular inflammation in response to treatment, suggesting utility of these approaches as outcome measures in future treatment trials in LVV.

The ease of implementation makes a qualitative strategy like PETVAS an attractive option for clinical assessment; however, there are some limitations in comparison to semi-quantitative approaches. Qualitative visual assessment requires reader experience and is subjective. Semi-quantitative approaches, while more time consuming and labor intensive, have superior reliability compared to PETVAS. The granularity and continuous scale of semi-quantitative scoring systems leads to better ability to discriminate change in PET activity across a wider range of values. Use of an ordinal scale like PETVAS with a ceiling limit of 27 may not capture important variability in patients with severe vascular inflammation, a situation where semi-quantitative metrics may be preferable or an opportunity to investigate improvements in qualitative scoring.

Semi-quantitative approaches correlated better than qualitative assessments with circulating markers of systemic inflammation; however, correlation was weak. Future biomarker discovery studies in LVV that use FDG-PET findings as a reference standard for disease activity should consider semi-quantitative metrics rather than quantitative metrics to quantify vascular inflammation for greater precision to detect candidate circulating biomarkers. In keeping with prior studies, the overall correlation of vascular inflammation with concentrations of acute phase reactants was poor (28).

TBRs and PETVAS achieved better performance characteristics than SUV when compared to reader interpretation of vascular PET activity, which is in line with a recent study by an independent group (30). TBR_{Liver} and TBR_{Blood} displayed near identical performance characteristics in association with clinical assessment of disease activity. However, TBR_{Liver} was more strongly associated with reader interpretation of vascular PET activity and with circulating acute phase reactants.

There are several study strengths to highlight. FDG-PET image acquisition and subsequent imaging interpretation was performed according to standardized protocols. Clinical and imaging assessments were performed independent of each other to enable unbiased comparisons. A prospective, longitudinal study design was employed, which is uncommon in vascular imaging studies in LVV but important to understand the utility of FDG-PET to detect change in vascular inflammation and to avoid bias inherent to retrospective study designs. Performance characteristics of PET assessment were tested against both reader interpretation of PET activity and physician assessment of clinical disease activity and performed well against both of these independent reference standards.

There are a few limitations to consider. This study was conducted in a single center using a specific imaging protocol, and these findings should be replicated in other cohorts. Specifically, the qualitative and quantitative imaging metrics reported here are a product of the methodology used for patient preparation, image acquisition, and image reconstruction at a single institution. Thus, the performance characteristics of discrete cutoffs for metrics, as applied in this study, will vary if applied broadly. This study compares the performance of different methods to measure

arterial FDG uptake as might be used in the clinical management of patients or in clinical trials of LVV. However, issues of feasibility and cost must be balanced against potential test utility.

CONCLUSION

Qualitative and semi-quantitative approaches to measure arterial FDG uptake are useful to detect and monitor vascular inflammation in LVV. Qualitative metrics, such as PETVAS, can be employed for FDG-PET assessment when simplicity and ease-of-interpretation are a priority, as is often the case in clinical practice or observational studies. Semi-quantitative metrics can be utilized for FDG-PET assessment when there is a need for greater precision, such as in randomized clinical trials or translational research focused on biomarker discovery.

KEY POINTS

Question: To compare the performance characteristics of qualitative versus semi-quantitative metrics of arterial FDG uptake to detect and monitor vascular inflammation by positron emission tomography.

Pertinent Findings: In this prospective, observational cohort study of 95 patients with large-vessel vasculitis, qualitative and semi-quantitative measurements of arterial FDG uptake were useful to monitor vascular inflammation.

Implications for Patient Care: Assessment of vascular inflammation by FDG-PET should be studied as an outcome measure in clinical trials of large-vessel vasculitis.

FINANCIAL DISCLAIMER

This work was supported through the intramural research program at the National Institute of Arthritis Musculoskeletal and Skin Disease (ZIA-AR-041199). The authors declare no conflicts of interest.

REFERENCES

1. Jennette JC, Falk RJ, Bacon PA, et al. 2012 revised International Chapel Hill Consensus Conference nomenclature of vasculitides. *Arthritis Rheum.* 2013;65:1-11.
2. Rosenbaum D, Millon A, Fayad ZA. Molecular imaging in atherosclerosis: FDG PET. *Curr Atheroscler Rep.* 2012;14:429-437.
3. Lee YH, Choi SJ, Ji JD, Song GG. Diagnostic accuracy of 18F-FDG PET or PET/CT for large vessel vasculitis. *Z Rheumatol.* 2016;75:924-931.
4. Prieto-González S, Depetris M, García-Martínez A, et al. Positron emission tomography assessment of large vessel inflammation in patients with newly diagnosed, biopsy-proven giant cell arteritis: a prospective, case–control study. *Ann Rheum Dis.* 2014;73:1388-1392.
5. Yamashita H, Kubota K, Mimori A. Clinical value of whole-body PET/CT in patients with active rheumatic diseases. *Arthritis Res Ther.* 2014;16:423.
6. Blockmans D, Ceuninck L de, Vanderschueren S, et al. Repetitive 18F-fluorodeoxyglucose positron emission tomography in giant cell arteritis: A prospective study of 35 patients. *Arthritis Care Res (Hoboken).* 2006;55:131-137.
7. Blockmans D, Bley T, Schmidt W. Imaging for large-vessel vasculitis. *Curr Opin Rheumatol.* 2009;21:19-28.
8. Both M, Ahmadi-Simab K, Reuter M, et al. MRI and FDG-PET in the assessment of inflammatory aortic arch syndrome in complicated courses of giant cell arteritis. *Ann Rheum Dis.* 2008;67:1030-1033.
9. Lee K-H, Cho A, Choi Y-J, et al. The role of 18F-fluorodeoxyglucose–positron emission tomography in the assessment of disease activity in patients with Takayasu arteritis.

- Arthritis Rheum.* 2012;64:866-875.
10. DeJaco C, Ramiro S, Duftner C, et al. EULAR recommendations for the use of imaging in large vessel vasculitis in clinical practice. *Ann Rheum Dis.* 2018;77:636.
 11. Slart RHJA, Glaudemans AWJM, et al. 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.
 12. Grayson PC, Tomasson G, Cuthbertson D, et al. Association of vascular physical examination findings and arteriographic lesions in large vessel vasculitis. *J Rheumatol.* 2012;39:303-309.
 13. Kermani TA, Diab S, Sreih AG, et al. Arterial lesions in giant cell arteritis: A longitudinal study. *Semin Arthritis Rheum.* 2019;48:707-713.
 14. Meller J, Strutz F, Siefker U, et al. Early diagnosis and follow-up of aortitis with [18F]FDG PET and MRI. *Eur J Nucl Med Mol Imaging.* 2003;30:730-736.
 15. Meignan M, Gallamini A, Meignan M, et al. Report on the first international workshop on interim-PET-scan in Lymphoma. *Leuk Lymphoma.* 2009;50:1257-1260.
 16. Hautzel H, Sander O, Heinzl A, et al. Assessment of large-vessel involvement in Giant Cell Arteritis with 18F-FDG PET: Introducing an ROC-Analysis–Based Cutoff Ratio. *J Nucl Med.* 2008;49:1107-1113.
 17. Rosenblum JS, Quinn KA, Rimland CA, et al. Clinical factors associated with time-specific distribution of 18F-fluorodeoxyglucose in large-vessel vasculitis. *Sci Rep.* 2019;9:15180.

18. Grayson PC, Alehashemi S, Bagheri AA, et al. 18F-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.
19. Rudd JHF, Myers KS, Bansilal S, et al. Atherosclerosis inflammation imaging with 18F-FDG PET: carotid, iliac, and femoral uptake reproducibility, quantification methods, and recommendations. *J Nucl Med*. 2008;49:871-878.
20. Zhuang H, Alavi A. 18-Fluorodeoxyglucose positron emission tomographic imaging in the detection and monitoring of infection and inflammation. *Semin Nucl Med*. 2002;32:47-59.
21. Arend WP, Michel BA, Bloch DA, et al. The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. *Arthritis Rheum*. 1990;33:1129-1134.
22. Hunder GG, Bloch DA, Michel BA, et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis Rheum*. 1990;33:1122-1128.
23. Langford CA, Cuthbertson D, Ytterberg SR, et al. A randomized, double-blind trial of abatacept (CTLA-4Ig) for the treatment of giant cell arteritis. *Arthritis Rheumatol*. 2017;69:837-845.
24. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-163.
25. Park SH, Goo JM, Jo C-H. Receiver operating characteristic (ROC) curve: Practical review for radiologists. *Korean J Radiol*. 2004;5:11-18.
26. Ripplinger J, Sullivan J. Does choice in model selection affect maximum likelihood analysis? *Syst Biol*. 2008;57:76-85.
27. Banerjee S, Quinn KA, Gribbons KB, et al. Effect of treatment on imaging, clinical, and

- serologic assessments of disease activity in large-vessel vasculitis. *J Rheumatol*. 2020;47:99-107.
28. Quinn KA, Ahlman MA, Malayeri AA, et al. Comparison of magnetic resonance angiography and ¹⁸F-fluorodeoxyglucose positron emission tomography in large-vessel vasculitis. *Ann Rheum Dis*. 2018;77:1165-1171.
29. Kerr GS, Hallahan CW, Giordano J, et al. Takayasu arteritis. *Ann Intern Med*. 1994;120:919-929.
30. Kang F, Han Q, Zhou X, et al. Performance of the PET vascular activity score (PETVAS) for qualitative and quantitative assessment of inflammatory activity in Takayasu's arteritis patients. *Eur J Nucl Med Mol Imaging*. 2020;47:3107-3117.
31. Arnaud L, Haroche J, Malek Z, et al. Is ¹⁸F-fluorodeoxyglucose positron emission tomography scanning a reliable way to assess disease activity in Takayasu arteritis? *Arthritis Rheum*. 2009;60:1193-1200.

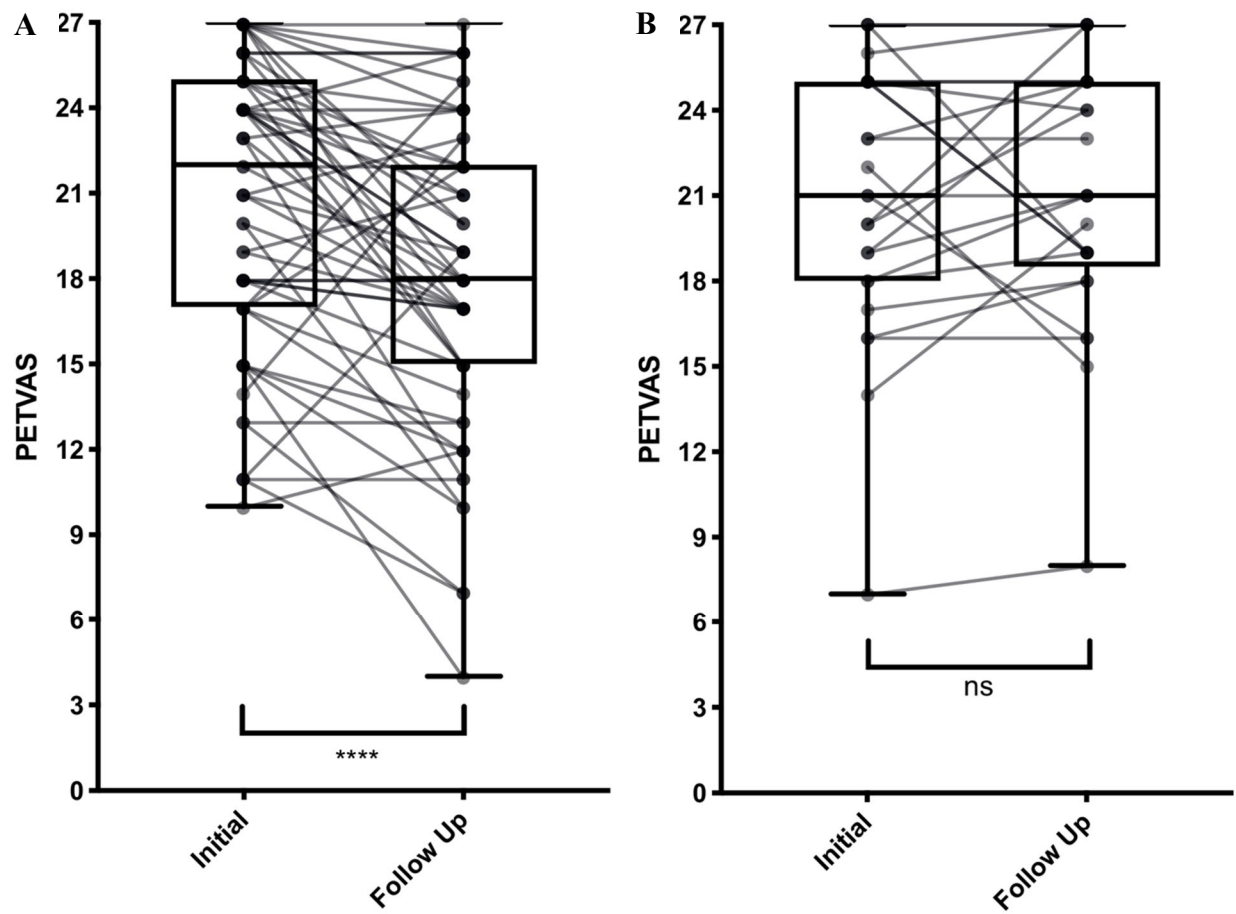


Figure 1. Response of Qualitative FDG-PET Assessment Scores to (A) Increased Pharmacological Treatment and (B) No Change to Pharmacological Treatment. Data is shown as box and whisker plots with paired comparisons. ‘ns’ = p-value > 0.05, ‘****’ = p-value < 0.001.

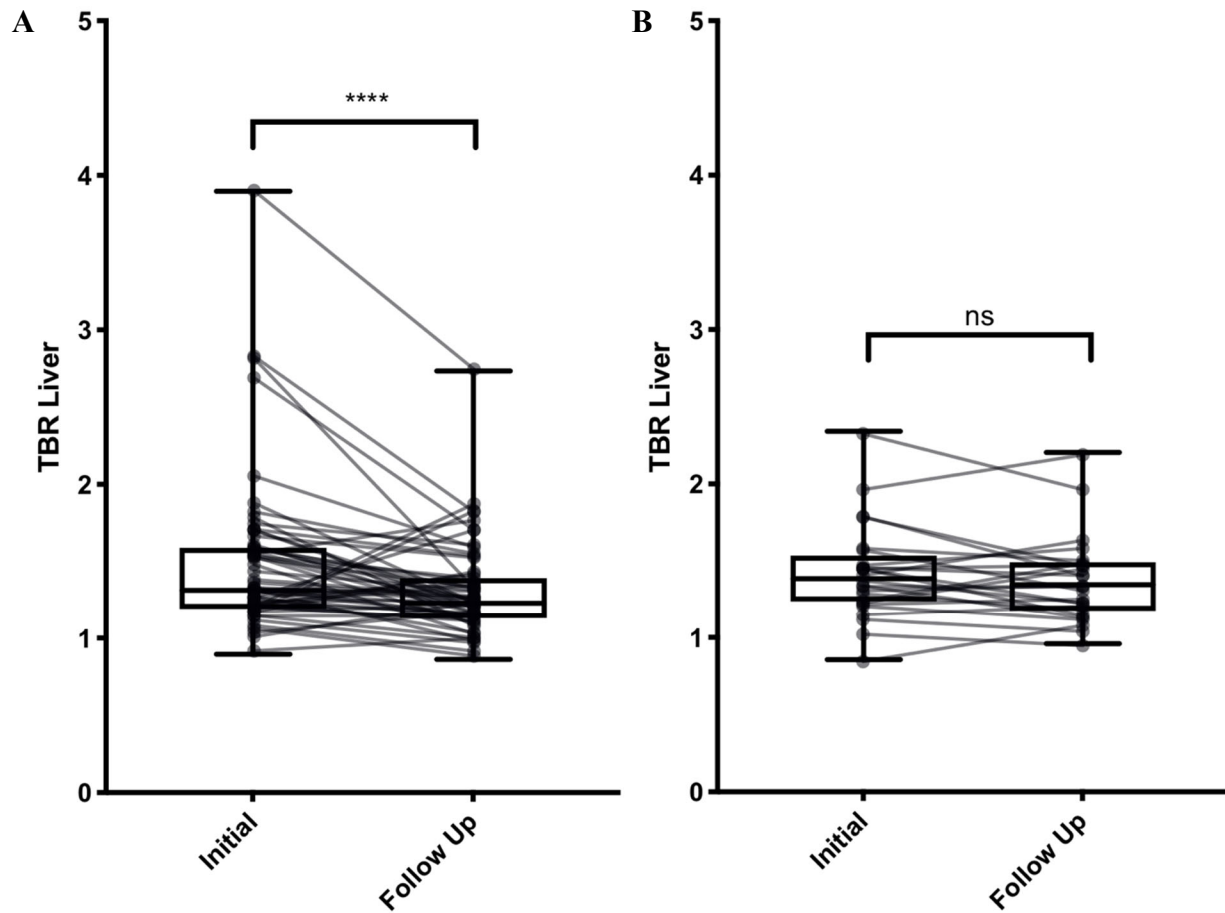


Figure 2. Response of Semi-Quantitative FDG-PET Assessment Scores to (A) Increased Pharmacological Treatment and (B) No Change to Pharmacological Treatment. Data is shown as box and whisker plots with paired comparisons. 'ns' = p-value > 0.05, '****' = p-value < 0.001.

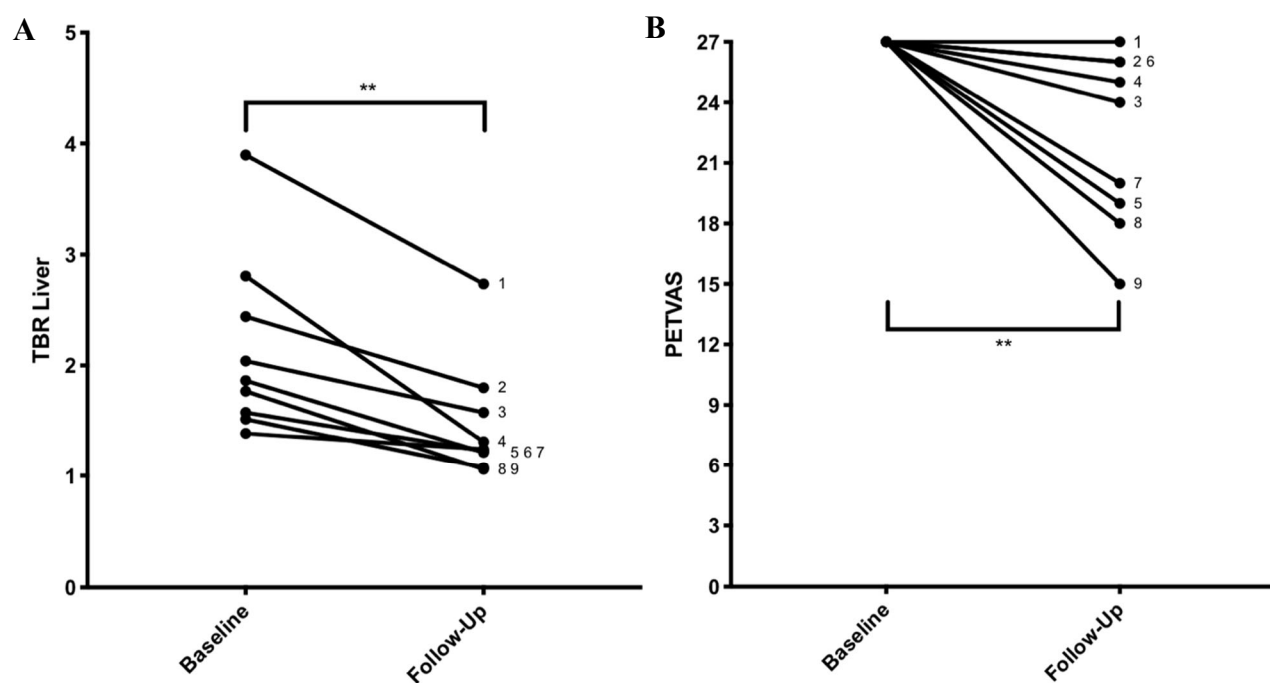


Figure 3. Change in FDG-PET Assessment Metrics in a Subset of Patients with Maximum Baseline PET Vascular Activity Scores using qualitative metrics (A) or semi-quantitative metrics (B). ‘ns’ = p-value > 0.05, ‘**’ = p-value < 0.01.

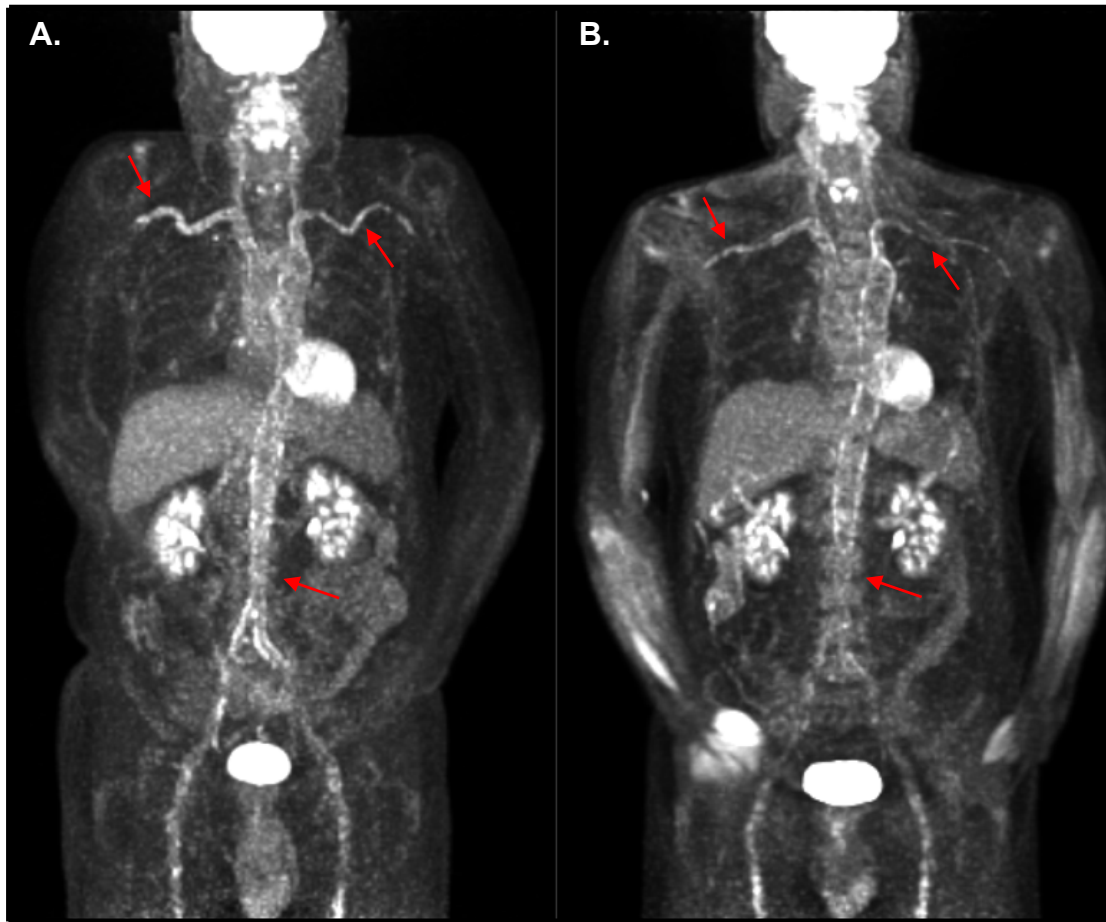


Figure 4. Improvement in FDG activity over time in a patient with severe vascular inflammation. (A) Baseline imaging study with PETVAS=27 (max score). (B) Follow-up imaging 6 months later with continued PETVAS=27 despite visual improvement in arterial FDG uptake and corresponding decrease in semi-quantitative metrics (TBR_{Liver}) from 3.90 to 2.73. Arrows show areas of increased FDG uptake in the right and left subclavian/axillary arteries and the abdominal aorta.

Tables

TABLE 1. Study Population Baseline Demographics

	Giant Cell Arteritis	Takayasu's Arteritis	Total
Patients (n)	52	43	95
PET-CT Studies Per Patient			
1 Study	22	19	41
2 Studies	11	14	25
≥ 3 Studies	19	10	29
Age (Years ± SD)	69.2 ± 8.9	34.3 ± 10.3	55.9 ± 19.4
Sex (Female, %)	41 (78.8)	33 (76.7)	74 (77.9)
BMI (± SD)	27.3 ± 5.5	25.9 ± 6.7	26.8 ± 6.0
CRP (± SD)	6.8 ± 14.1	10.7 ± 21.7	8.3 ± 17.4
ESR (± SD)	17.3 ± 19.1	19.4 ± 16.6	18.1 ± 18.2
Daily Prednisone (mg ± SD)	9.0 ± 13.9	7.2 ± 11.3	8.33 ± 12.9
Disease Duration (Years ± SD)	2.98 ± 2.41	10.9 ± 10.7	6.06 ± 7.94

n = number; SD = standard deviation; BMI = body mass index; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate

TABLE 2. Performance Characteristics of Semi-Quantitative and Qualitative PET Assessment Methods in Association with Subjective Reader Interpretation of PET Activity

	Descriptive Statistics				GLMM Evaluation	
	PET Active (Mean \pm SEM)	PET Inactive (Mean \pm SEM)	AUC (95% CI)	Optimal Cut-off	P Value	AIC
LVV (N)	147	62				
SUV _{Artery}	3.41 \pm 0.10	2.78 \pm 0.08	0.67 (0.60 - 0.75)	3.36	<0.0001	226.9
TBR _{Liver}	1.45 \pm 0.04	1.09 \pm 0.02	0.85 (0.80 - 0.90)	1.22	<0.0001	179.0
TBR _{Blood}	2.59 \pm 0.07	2.03 \pm 0.03	0.80 (0.74 - 0.86)	2.31	<0.0001	198.2
PETVAS	20.83 \pm 0.41	12.87 \pm 0.58	0.87 (0.83 - 0.92)	19.5	<0.0001	165.2

LVV = large vessel vasculitis; N = number; SEM = standard error of the mean; AUC = area under curve; CI = confidence interval; AIC = Akaike information criterion

TABLE 3. Performance Characteristics of Semi-Quantitative and Qualitative PET Assessment Methods in Association with Physician Assessment of Clinical Disease Activity

	Descriptive Statistics				GLMM Evaluation	
	Clinical Active (Mean \pm SEM)	Clinical Remission (Mean \pm SEM)	AUC (95% CI)	Optimal Cut-off	P Value	AIC
LVV (N)	75	131				
SUV _{Artery}	3.43 \pm 0.13	3.11 \pm 0.09	0.59 (0.51 - 0.68)	3.58	0.0293	255.2
TBR _{Liver}	1.46 \pm 0.06	1.27 \pm 0.03	0.66 (0.58 - 0.73)	1.46	0.0002	253.1
TBR _{Blood}	2.60 \pm 0.09	2.31 \pm 0.06	0.65 (0.57 - 0.73)	2.39	0.0003	254.7
PETVAS	20.6 \pm 0.56	17.3 \pm 0.55	0.65 (0.57 - 0.73)	22.5	0.0004	239.5

LVV = large vessel vasculitis; N = number; SEM = standard error of the mean; AUC = area under curve; CI = confidence interval; AIC = Akaike information criterion

TABLE 4. Correlation of Semi-Quantitative and Qualitative PET Assessment Methods to Acute Phase Reactants (CRP and ESR)

	CRP		ESR	
	Spearman r	P Value	Spearman r	P Value
SUV _{Artery}	0.19	< 0.01	0.14	0.04
TBR _{Liver}	0.20	< 0.01	0.15	0.03
TBR _{Blood}	0.11	0.14	0.02	0.78
PETVAS	0.08	0.29	0.08	0.27

CRP = C-reactive protein; ESR = erythrocyte sedimentation rate

Graphical Abstract