

## Article

# Performance of Deauville Criteria in $^{18}\text{F}$ FDG-PET/CT Diagnostics of Giant Cell Arteritis

Jeffrey Siefert <sup>1,\*</sup>,†, Jonas Kaufmann <sup>1,†</sup>, Felix Thiele <sup>1</sup>, Thula Walter-Rittel <sup>2</sup>, Julian Rogasch <sup>1,3</sup>, Robert Biesen <sup>4</sup>, Gerd R. Burmester <sup>4</sup>, Holger Amthauer <sup>1</sup>, Udo Schneider <sup>4</sup> and Christian Furth <sup>1</sup>

<sup>1</sup> Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Nuclear Medicine, Augustenburger Platz 1, 13353 Berlin, Germany

<sup>2</sup> Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Radiology, Augustenburger Platz 1, 13353 Berlin, Germany

<sup>3</sup> Berlin Institute of Health at Charité—Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany

<sup>4</sup> Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Rheumatology and Clinical Immunology, Charitéplatz 1, 10117 Berlin, Germany

\* Correspondence: jeffrey.siefert@charite.de

† These authors contributed equally to this work.

**Abstract:** In this retrospective study, PET/CT data from 59 patients with suspected giant cell arteritis (GCA) were reviewed using the Deauville criteria to determine an optimal cut-off between PET positivity and negativity. Seventeen standardised vascular regions were analysed per patient by three investigators blinded to clinical information. Statistical analysis included ROC curves with areas under the curve (AUC), Cohen's and Fleiss' kappa ( $\kappa$ ) to calculate sensitivity, specificity, accuracy, and agreement. According to final clinician's diagnosis and the revised 2017 ACR criteria GCA was confirmed in 29 of 59 (49.2%) patients. With a diagnostic cut-off  $\geq 4$  (highest tracer uptake of a vessel wall exceeds liver uptake) for PET positivity, all investigators achieved high accuracy (range, 89.8–93.2%) and AUC (range, 0.94–0.97). Sensitivity and specificity ranged from 89.7–96.6% and 83.3–96.7%, respectively. Agreement between the three investigators suggested 'almost perfect agreement' (Fleiss'  $\kappa = 0.84$ ). A Deauville score of  $\geq 4$  as threshold for PET positivity yielded excellent results with high accuracy and almost perfect inter-rater agreement, suggesting a standardized, reproducible, and reliable score in diagnosing GCA. However, the small sample size and reference standard could lead to biases. Therefore, verification in a multicentre study with a larger patient cohort and prospective setting is needed.

**Keywords:**  $^{18}\text{F}$ FDG-PET/CT; giant cell arteritis; deauville criteria; inter-rater agreement



**Citation:** Siefert, J.; Kaufmann, J.; Thiele, F.; Walter-Rittel, T.; Rogasch, J.; Biesen, R.; Burmester, G.R.; Amthauer, H.; Schneider, U.; Furth, C. Performance of Deauville Criteria in  $^{18}\text{F}$ FDG-PET/CT Diagnostics of Giant Cell Arteritis. *Diagnostics* **2023**, *13*, 157. <https://doi.org/10.3390/diagnostics13010157>

Academic Editors: Andor W.J.M. Glaudemans and Riemer H.J.A. Slart

Received: 27 November 2022

Revised: 28 December 2022

Accepted: 29 December 2022

Published: 3 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Combined Positron-Emission-Tomography/Computerized-Tomography (PET/CT) using  $^{18}\text{F}$ Fluorodeoxyglucose ( $^{18}\text{F}$ FDG) is an established molecular imaging modality to visualize glucose metabolism [1]. Due to the accumulation and trapping of  $^{18}\text{F}$ FDG within activated inflammatory cells [2], it has been used to detect inflammation in infectious and inflammatory diseases [3]. Thus, PET-based imaging of patients suffering from giant cell arteritis (GCA) was performed as early as 1999 [4]. GCA is increasingly recognized as a systemic disease, which can involve discontinuous vascular regions of the entire body [5]. In absence of temporal involvement diagnosis can be difficult since typical symptoms of cranial involvement are commonly absent in extra cranial disease [6]. In addition, other typical symptoms such as fever, myalgia, weight loss and elevation of blood inflammatory markers are unspecific [7].

$^{18}\text{F}$ FDG-PET/CT has proven to be sensitive in the diagnosis of GCA [8] and may in some cases be the only modality suitable for diagnosis [9]. However, there is still a lack

of standardization in interpretation of  $^{18}\text{F}$ FDG-PET/CT scans in patients suffering from GCA [10]. In particular, a robust, reliable, and endpoint-oriented definition of PET-positive and PET-negative has yet to be established. Thus, the data in the literature published to date are only comparable to a very limited extent; thus, limiting the widespread use of  $^{18}\text{F}$ FDG-PET/CT in patients suffering from GCA.

In comparison, assessment of disease activity in patients suffering from lymphoma has been successfully standardized using the Deauville criteria [11], with observations showing high degree of accuracy and inter-rater reliability [12,13]. In the Deauville system, up to 5 points are scored for the intensity of FDG enhancement. By comparing FDG enhancement in the region of interest with that of mediastinum and liver, an internal standardisation is provided.

Thus, we are proposing the Deauville criteria as a new way of diagnosing patients suffering from GCA. The applicability was evaluated by multiple readers using the criteria in patients versus a control group and assessing multiple vascular regions to evaluate robustness and reliability of the Deauville criteria. In addition, we investigated whether the applied score also allows a distinct differentiation between PET-positive versus PET-negative.

## 2. Materials and Methods

### 2.1. Study Population

This monocentric, cross-sectional study screened all patients for study inclusion who received  $^{18}\text{F}$ FDG-PET/CT imaging in the nuclear medicine department of Charité–Universitätsmedizin Berlin from April 1st, 2016, to June 28th, 2020. Inclusion criteria were age >40 years, prednisolone and prednisolone equivalent doses <20 mg/d at the time of PET imaging, glucocorticoids for a duration <72 h before PET and no vascular implants). Standard of reference for patients with active disease were the proposed 2017 expansion of the ACR criteria for GCA [14]. Additionally, all patients needed to have a documented indication for initiation, escalation or change of therapy at the time of the scan, as per obtainable records. The clinical diagnosis incorporated the patients' medical history, physical examination, clinical symptoms, laboratory tests and imaging, including PET/CT studies examined in this study. Clinicians did not receive reports based on the evaluated Deauville criteria at the time of initial diagnosis, as it was not implemented as clinical standard.

Out of 234 patients with suspected disease, the data sets of 29 patients with confirmed GCA were eligible for analysis. The patient group was matched with a control group by random selection from the patient pool with negative final diagnosis, who fulfilled the same in- and exclusion criteria, thus resulting in a patient population of 59 patients (Table 1).

**Table 1.** Description patient cohort.

Variable	Patient Cohort	Control Cohort
Total (n)	29	30
Female (n)	16/29	18/30
Male (n)	13/29	12/30
Age (years)	66 (range, 51–83)	66 (range, 49–79)
Activity (MBq)	269 (IQR, 245–302.5)	274.5 (IQR, 251–300.25)
Uptake time (min)	67 (IQR, 64–80)	71.5 (IQR, 65–84)
Blood glucose (mmol/L)	5.4 (IQR, 4.9–6)	5.6 (IQR, 5–6.3)
CRP (mg/L)	36.9 (IQR, 18.95–65.1)	38.15 (IQR, 3.8–84.9) *

Description of analysed cohort, in numbers for total patients/control patients in the examined cohort, as well as ratio female to male. Age, administered activity, uptake time between administration of  $^{18}\text{F}$ FDG-PET/CT imaging as well as blood glucose are given as median and respective IQR. 'n', count; MBq, Megabecquerel; IQR, interquartile range. \* CRP not documented in n = 1 case.

All analyses were in agreement with the Declaration of Helsinki and the ethical standards of the research ethics committee of the Charité—Universitätsmedizin Berlin (ID number: EA1/182/20; approved at 30th September 2020).

## 2.2. PET/CT Imaging

The tracer  $^{18\text{F}}$ FDG was used for PET imaging, and scans were performed on two PET/CT devices. Patients were advised to fast for at least 6 hrs before examination. Scans were conducted with acquisition times of 2 to 3 min per bed position and had to include at least the base of the skull to the proximal femurs.

Forty-two of the 59 patients received a scan on a digital PET device (General Electric<sup>®</sup> Healthcare, Discovery MI, Chicago, Illinois, United States; silicon photomultipliers [SiPM], 3-ring detector setting, Time of flight [TOF] capability, system sensitivity of 7.3 cps/kBq). Median injected activity equalled 279.5 MBq (range, 227–397 MBq) with a median uptake time of 67.5 min (range, 58–111 min). PET raw data were reconstructed iteratively with Bayesian penalized likelihood reconstruction (GE “Q.Clear”) with a penalization factor  $\beta$  of 450 (matrix,  $256 \times 256$ ; voxel size,  $2.73 \times 2.73 \times 2.78 \text{ mm}^3$ ).

In 17/59 patients, scans were performed on a non-digital PET -device with photomultiplier tubes (Philips<sup>®</sup>, Gemini TF16, Eindhoven, The Netherlands; lutetium-yttrium oxyorthosilicate scintillation crystals, TOF technology, system sensitivity of 6.6 cps/kBq). Median injected activity equalled 252 MBq (range, 228–300 MBq), median uptake time was 72 min (range, 58–120 min). PET raw data were reconstructed using iterative reconstruction (ordered subset expectation maximization; OSEM) with TOF analysis (BLOB-OS-TF; iterations, 3; subsets, 33; filter, ‘smooth’ [kernel width, 14.1 cm; relaxation parameter, 0.7]; matrix,  $144 \times 144$ ; voxel size,  $4.0 \times 4.0 \times 4.0 \text{ mm}^3$ ).

With both PET scanners, non-enhanced low-dose CT data served for attenuation correction. Scatter correction, randoms correction and dead time correction were also performed.

Values for injected activity, uptake time and blood glucose separated by diagnosis can be found in Table 1.

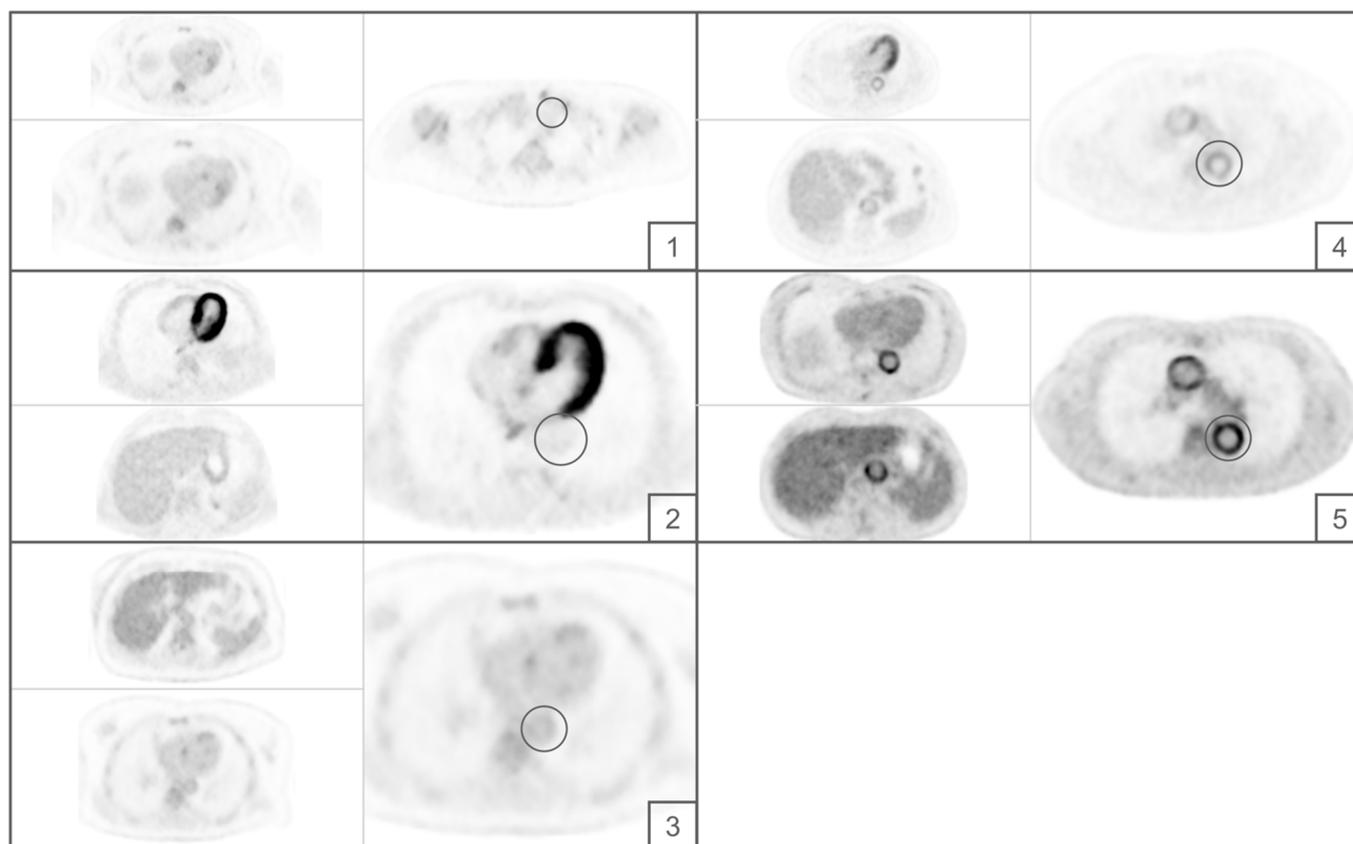
## 2.3. Image Assessment

Imaging data were assessed independently by three readers with varying level of experience (reader1 [R1], >10 years; reader2 [R2], >3 years; reader3 [R3], <6 months). Seventeen vascular regions (aorta with its main branches; ascending, thoracic, and abdominal aorta plus aortic arch; assessed branches included the innominate artery as well as carotid, vertebral, subclavian, axillary, iliac and femoral arteries) were scored individually per patient. When the arterial regions were not fully displayed (e.g., femoral arteries) the visible areas were scored. The highest scored lesion was decisive for the patient-based analysis. The readers were blinded to the clinical information and the diagnosis of the patients.

Readers were trained in the application of the Deauville criteria (Figure 1) by reference training data sets not included in the final analysis.

Data sets were analysed visually with the axial PET images. However, readers were asked to measure exemplary standardized uptake values (SUV<sub>mean</sub>) of the respective vessel segment using 2D-ROIs after assigning the Deauville scores. The points of measurement were chosen by the readers as per highest visual uptake in each segment.

Furthermore, each reader had to perform a standardised measurement of the SUV<sub>mean</sub> in the liver (right liver lobe) and in the right cardiac atrium (blood pool). The liver and cardiac volumes of interest (VOI) should only include areas with physiological tracer uptake; in particular, in the area of the right atrium, parts of the myocardium should not be included by the VOI. The SUV-values were measured for internal quality control, and to correct for visual phenomena. Fusion data sets should only be used for anatomical validation of the respective vessel segment. No specific pre-sets for window settings were used.



**Figure 1.** Visual representation of the Deauville criteria. Numbers 1 to 5 represent the Deauville score for each marked vessel. 1 equals no uptake. 2 equals uptake less than the mediastinal blood pool. 3 equals uptake higher than uptake in the mediastinal blood pool but lower than liver uptake. 4 equals uptake higher than liver uptake. 5 equals uptake distinctly higher than liver uptake.

#### 2.4. Statistical Analysis

Computations were primarily done using SPSS 27 (IBM Corporation, Armonk, NY, United States). Histograms, Q-Q plots, and Shapiro–Wilk-Test concluded a non-parametric distribution of data. Unless otherwise mentioned, descriptive results are presented as median, interquartile range (IQR), and range. Further calculations for receiver operating characteristics (ROC) with areas under the curve (AUC) and their 95%-confidence interval (95%-CI) were performed. Optimal cut-off values were determined based on the minimal distance  $d$  to the point (0, 1), defined as follows:

$$\sqrt{(1 - \text{Sensitivity})^2 + (1 - \text{Specificity})^2}$$

Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy for cut-off points were calculated using commonly known equations, while confidence intervals were calculated with an excel-tool from “ACOMED Statistik” [Dr. Thomas Keller, Leipzig, [www.acomed-statistik.de](http://www.acomed-statistik.de) (accessed on 19 December 2021)].

To compare two readers regarding their accuracy, McNemar’s test for paired nominal data was used.

Furthermore, variability between paired readers in evaluating each vascular region as well for diagnosing patients was computed using Cohen’s Kappa ( $\kappa$ ). For investigating overall agreement, Fleiss’  $\kappa$  was calculated. Computations were done by classifying the highest Deauville-based assessment of all vascular regions as positive or negative for GCA, as defined by a chosen threshold. To classify and interpret  $\kappa$ -values, the benchmarks by Landis and Koch [15] were used. Confidence intervals for  $\kappa$  were calculated like

McHugh [16] proposed, in order to compare confidence intervals to investigate significant difference between the pairs of reviewers. Statistical significance was assumed at  $\alpha = 0.05$ .

### 3. Results

#### 3.1. Patient Characteristics

In the elected time frame, 234 patients with suspected vasculitis received PET/CT imaging. 35 of whom fulfilled the revised 2017 criteria for GCA. Because of glucocorticoid treatment above the limit ( $n = 2$ ), image quality ( $n = 2$ ), myocarditis ( $n = 1$ ) and vascular implants ( $n = 1$ ) patients had to be excluded resulting in a total of 29 patients with GCA. The patient with myocarditis was excluded due to possible affection of the ascending aorta. In 25 of the 29 patients, the initial diagnosis was made during the inpatient workup. In 4 cases, PET/CT was performed due to a relapse.

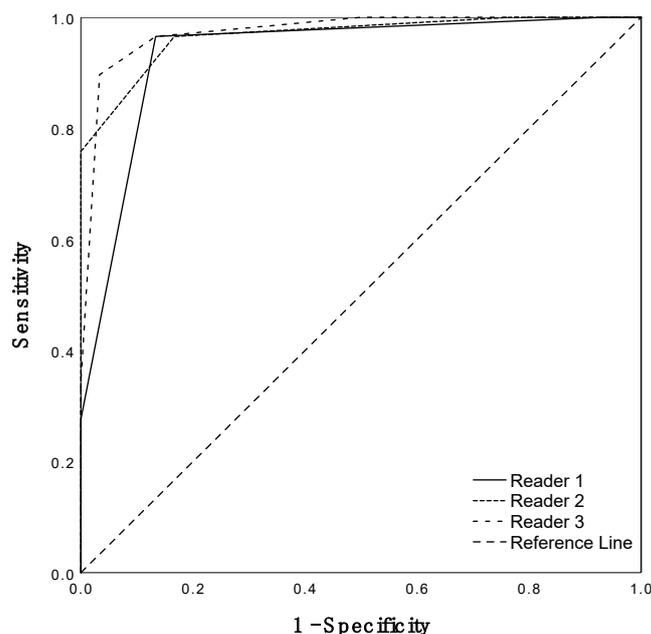
11/29 patients showed purely nonspecific symptoms, such as fatigue, fever, night sweats or weight loss. Ten others presented with joint or muscle pain in addition to non-specific symptoms. Two patients showed myalgia or arthralgia. Only 4/29 patients exhibited specific symptoms of cranial GCA, such as jaw claudication or loss of visual acuity. Twelve patients exhibited unspecific cranial symptoms such as headaches or dysphagia, with two patients being diagnosed after apoplexy workup. One patient was described as oligosymptomatic. In the control cohort, the most often described symptoms were myalgia or arthralgia (20 patients), overlapping with non-specific symptoms (fatigue, fever, night sweats) in 19 patients. Final diagnoses of the control group were rheumatoid arthritis ( $n = 8$ ), connective tissue diseases ( $n = 4$ ), idiopathic serositis ( $n = 3$ ), inflammation of unknown origin ( $n = 3$ ), gout arthritis ( $n = 2$ ), sarcoidosis ( $n = 2$ ), gastrointestinal disease ( $n = 2$ ), malignancy ( $n = 1$ ), reactive arthritis ( $n = 1$ ), Adult-onset Still's disease ( $n = 1$ ), polymyalgia rheumatica ( $n = 1$ ) ANCA vasculitis ( $n = 1$ ) and hypertensive crisis ( $n = 1$ ).

In addition to PET/CT, other imaging studies performed during clinical workup were (20/59), MRI (13/59), or CT (7/59). Temporal artery biopsy was performed in 2/29 patients, with confirmation in one patient. 22/59 patients did not receive any relevant vascular imaging.

#### 3.2. Diagnostic Performance

In a per patient analysis, R1 scored patients with a Deauville score of 5 ( $n = 8$ ), 4 ( $n = 24$ ), 3 ( $n = 25$ ), 2 ( $n = 2$ ), and zero patients with a maximum score of 1. R2 scored 5 ( $n = 22$ ), 4 ( $n = 11$ ), 3 ( $n = 19$ ), 2 ( $n = 7$ ), as well no patient with a score of 1. R3 differed with 5 ( $n = 10$ ), 4 ( $n = 17$ ), 3 ( $n = 5$ ), 2 ( $n = 12$ ) and 1 ( $n = 15$ ).

Leading lesion-based ROC analysis revealed that a Deauville score of 4 or higher was clinically the most accurate in distinct differentiation of patients from controls (Figure 2). Highest diagnostic accuracy was achieved by R3 with 93.2%, resulting in a sensitivity and specificity of 89.7% and 96.7% (PPV, 96.3%; NPV, 90.6%), with an AUC of 0.97 (95%-CI, 0.94–1.0). Results for R1 and R2 were similar, but concluded higher sensitivities, while accuracy and specificity were not as high as in R3. Exact values, results on PPV, NPV and respective 95%-CIs can be found in Table 2. R2 and R3 differed significantly in accuracy ( $p < 0.05$ ), while the pairs with R1 (R1/R2 and R1/R3) showed no significant differences in accurate diagnosis ( $p > 0.05$ ).



**Figure 2.** Patient-based ROC analysis per reader ROC curves in a per-patient analysis for all three readers. Calculation based on the highest Deauville score assignment for any vessel per patient.

**Table 2.** Diagnostic performance and ROC analysis.

Reader	TP	TN	FP	FN	Accuracy		Sensitivity		Specificity		PPV		NPV		AUC	
					%	95%-CI	%	95%-CI	%	95%-CI	%	95%-CI	%	95%-CI	%	95%-CI
1	28	26	4	1	91.5	81.3–97.2	96.6	82.2–99.9	86.7	69.3–96.2	87.5	71.0–96.5	96.3	81.0–99.9	0.94	0.87–1.00
2	28	25	5	1	89.8	79.2–96.2	96.6	82.2–99.9	83.3	65.3–94.4	84.9	68.1–94.9	96.2	80.4–99.9	0.97	0.93–1.01
3	26	29	1	3	93.2	83.5–98.1	89.7	72.7–97.8	96.7	82.8–99.9	96.3	81.0–99.9	90.6	75.0–98.0	0.97	0.94–1.01
Average					91.5		94.3		88.9		89.6		94.3		0.96	

TP, true positives; TN, true negatives; FP, false positives; FN, false negatives; PPV, positive predictive values; NPV, negative predictive values; AUC, area under the curve from receiver operating characteristics; 95%-CI, 95%-confidence interval.

In a case-by-case analysis, no patterns in false positive or false negative findings emerged.

### 3.3. Comparison with Cut-Off $\geq 3$

While a cut-off of  $\geq 4$  resulted in best clinical outcome in the patient cohort, results for an alternative cut-off  $\geq 3$  were also analysed. In R3, accuracy with cut-off  $\geq 3$  remained similar with no significant difference ( $p > 0.05$ ) at 91.5% (95%-CI, 81.3–97.2%), but sensitivity increased from 89.7% (95%-CI: 83.5–98.1%) to 96.7% (95%-CI: 82.2–99.9%) and specificity decreased from 96.7% (95%-CI: 82.8–99.9%) to 86.7% (95%-CI: 69.3–96.2%). For R1 and R2 accuracy decreased significantly ( $p < 0.05$ ) from 91.5% (95%-CI: 81.3–97.2%) to 52.5% (95%-CI, 39.1–65.7%) and from 89.8% (95%-CI: 79.2–96.2%) to 61.0% (95%-CI: 47.4–73.5%). Sensitivity was 100% (95%-CI: 88.1–100%) for both readers while specificity decreased to 6.7% (95%-CI: 0.8–22.1%) and 23.3% (95%-CI: 9.9–42.3%).

### 3.4. Inter-Rater Agreement

When comparing the highest scoring vascular regions, agreement was insufficient. However, in dichotomizing the criteria in PET positive ( $\geq 4$ ) and negative (1–3), readers achieved substantial to almost perfect agreement [15].

Most concordant ratings were achieved by the pair R1/R2 with 56/59 (94.9%) matching reads, resulting in  $\kappa = 0.90$  [95%-CI, 0.76–1.0 (almost perfect agreement)]. R1/R3 diagnosed 54/59 (91.5%) patients equally with  $\kappa = 0.83$  [95%-CI: 0.65–1.0 (almost perfect agreement)]. The pair R2/R3 achieved 53/59 (89.8%) equal reads,  $\kappa = 0.80$  [95%-CI, 0.64–0.96 (substantial

agreement)]. Because of overlap in confidence intervals, no significant difference between the pairs of readers can be assumed.

In 52 subjects (88.1%), concordant Deauville-based diagnoses were observed between all three readers, two of whom were concordant between readers but deviated from the reference diagnosis. Overall, resulting in Fleiss'  $\kappa$  of 0.84 [95%-CI, 0.69–0.99 (almost perfect)].

#### 4. Discussion

GCA is the most common of the vasculitides [17] and, if not detected and treated in a timely and properly manner will become a devastating disease by causing permanent damage. The diagnostic delay for GCA is prolonged if cranial symptoms are absent [18], due to the unspecific symptoms in which the disease manifests. In these cases,  $^{18}\text{F}$ FDG-PET/CT has shown potential as a valuable diagnostic tool [10,19]. However, the widespread use of  $^{18}\text{F}$ FDG-PET/CT remains limited by the lack of standardization of image interpretation and the absence of reliable and reproducible diagnostic criteria [10,20,21].

Therefore, we aimed to evaluate a standardized, PET-based approach for diagnosis of GCA—the *Deauville criteria* [11]. Choosing this score was based on its clear and concise definition of  $^{18}\text{F}$ FDG uptake patterns with reference to areas of physiological uptake as well as its proven high interobserver reliability for patients suffering from Hodgkin- and non-Hodgkin lymphomas [13,22,23]. Hence, leading to its adaptation for staging and response assessment in prospective studies in lymphoma patients [24–26] and ultimately to the use of  $^{18}\text{F}$ FDG-PET/CT as standard of care in clinical routine [27]. Thus, we evaluated whether this dedicated score is robustly and reliably applicable in patients with GCA. In addition, the clinical relevance of possible thresholds in to differentiate PET-negative versus PET-positive results was analysed.

We found that this standardized approach yielded excellent results in diagnostic performance with an average accuracy of 91.5%, sensitivity of 94.3% and specificity of 88.9%. In comparison, two meta-analyses from Besson et al. in 2011 [9] and Lee et al. in 2016 [28] reported inferior results to our study with sensitivities of 80% and 83.3% and similar specificities of 89% and 89.6%. However, contrary to our results, previous studies reported higher specificity than sensitivity [29–32]. In the context of everyday clinical practice, the higher sensitivity achieved in the current analysis could improve the therapeutic process and avoid under-treatment.

Further analyses concluded an overall agreement of Fleiss'  $\kappa$  of 0.84, which corresponds to an almost perfect agreement [15]. Despite the different levels of experience, no significant differences between the pairs of reviewers were detected, which we attribute to the standardized approach. Differences in measurement between the readers can largely be explained by methodical deficits of the Deauville criteria, as scores of 4 and 5 are not clearly defined, as well as variance in the leading lesions chosen by visual analysis. Compared to magnetic resonance angiography [33] with a Cohen's  $\kappa = 0.73$ , agreement seems to be favourable in PET/CT, concordant with the results of Meller et al. who already demonstrated superiority of PET over MRI in early stages of inflammation in 2003 [34].

As early as 2015, Lensen et al. [35] reported superb agreement for extra cranial GCA in  $^{18}\text{F}$ FDG-PET/CT with a  $\kappa = 0.96$  for 4 readers, using an uptake higher than liver uptake as visual threshold. The average sensitivity and specificity were 100% and 98%, respectively. Other tested methods were diagnosis by first impression as well as differing visual thresholds, namely, higher than the femoral artery uptake, or equal to the liver uptake. These approaches, however, led to inferior results, strengthening the argument for a visual threshold higher than the liver uptake. However, these results are weakened as the reported outcome was calculated after a consensus meeting, prior to which up to 1/3 of diagnoses were discordant [35]. The prior consensus meeting and the difference between the two examiners with least experience (two years compared to less than six months in our study) suggest that the aforementioned study is less reflective of a clinical real world setting.

Although PET/CT has been playing an important role in GCA diagnostics for many years, no real standardization and no definite cut-off between PET-positive and PET-negative could be established, and even most recent recommendations (joint recommendation of the EANM, SNMMI and PIG) remain indefinite [10]. For that reason, numerous systems have been developed, but neither semi-quantitative nor visual grading systems using either liver or blood pool activity as reference have been able to establish themselves [36]. In contrast to a proposed 4-point scale where uptake at liver level is classified as possibly PET-positive, the Deauville criteria only differentiate between uptake higher or lower than the liver uptake. For vessel wall lesions located close to the physiological reference regions (e.g., liver, and mediastinal blood pool), comparison of 'equal to liver uptake' may be possible. However, with larger distances between lesion and reference structures, background uptake perceptual phenomena, like the Mach effect [37], become relevant, making grading of 'equal to liver' almost impossible. As this phenomenon is also known from the validation phase of the Deauville criteria for the use of  $^{18}\text{F}$ FDG-PET in patients with lymphoma [38,39] it is known that it can be partially circumvented by giving the liver uptake a semi-quantitative value, which then corresponds to a punctual threshold—the standardized uptake value.

This is a possible explanation of the high interobserver agreement in our study, and why no consensus read was necessary. Overall, this perceptual phenomenon may explain the slight advantage in accuracy in our analysis and Lensen et al. (2015) [35] in comparison to the meta-analyses by Besson et al. (2011) and Lee et al. (2016) [9,28].

Our study deals with several limitations which must be considered. Most importantly, the small number of patients included in the study. Due to the rarity of the disease and the unspecific symptoms, only a small patient collective with confirmed GCA has emerged from the much larger collective with suspicion of GCA. Another factor in the small sample size are the in- and exclusion criteria, which were based on the conditions of usefulness for a PET examination. However, the final patient cohort and control cohort were not significantly different in CRP, further indicating the robustness and reliability of our conclusions.

Due to the retrospective design, incubation times were not standardized. Studies have shown, that prolonged uptake times lead to better target to background ratios in assessment of vascular pathologies [40,41]. Since the liver has shown to be the superior reference structure in this study, we do not expect a significant influence of the varying uptake times on the results.

In addition, selection bias cannot be excluded, as  $^{18}\text{F}$ FDG-PET/CT is not the first line imaging modality and is only recommended when no other imaging, such as color-coded ultrasound, leads to conclusive results [42]. Therefore, most patients in our analysis will not have shown typical symptoms of GCA, thus the 1990 ACR criteria were not appropriate for our population as they are limited to cranial involvement [43].

Due to the aforementioned reasons, and due to the resolution limits of the scanners used the cranial vessels were not evaluated. An extension of the 1990 classification criteria has been proposed [14] reflecting that the inclusion of imaging modalities has reached clinical routine years ago but may differ depending on local availability. Since the extended criteria have already been used in large clinical trials [44], we elected to choose the 2017 revised criteria as our standard of reference. However, the chosen reference standard is partially based on imaging data, including  $^{18}\text{F}$ FDG-PET/CT, which reduces the general transferability of the results.

As the Deauville criteria differentiate five scores in metabolic activity, its application may not be limited to discrimination between PET-positive versus PET-negative results at the time of baseline imaging but may also have predictive value. Huang et al. (2021) [45] have demonstrated a prognostic value of a Deauville score of 5 for true disease progression at end of treatment in extra-nodal natural killer/T-cell lymphoma. Therefore, discrimination of scores 4 and 5 as proposed by Barrington et al. (2014) [46] may prove its value in

differentiating patients at higher risk of vascular complications or relapse. To pursue this issue, further investigations are needed with larger sample sizes in a prospective setting.

Furthermore, a to be established validated standardized PET-reading could serve as a read-out parameter for successful therapy in GCA in clinical studies as well as in clinical routine.

## 5. Conclusions

The Deauville criteria, known from lymphoma assessment, appear to be applicable also in patients with GCA. The score is robust and reliably applicable. Diagnostic accuracy to detect patients with GCA was similarly high at different levels of experience. Vessel wall uptake with a Deauville score of  $\geq 4$  (i.e., higher than the physiological uptake of the liver) indicated disease in almost all cases. Because of this standardized evaluation approach, the application within clinical trials using a multicentre setting appears attractive. To what extent the results shown are also transferable to the situation of response assessment cannot be answered by the present data and requires additional analysis in further, preferably prospective, studies.

**Author Contributions:** Conceptualization, C.F. and H.A.; Data curation, J.K. and J.S.; Formal analysis, J.S. and J.R.; Investigation: J.K., F.T., T.W.-R. and C.F.; Methodology, J.K., J.S. and J.R.; Project administration, J.S. and C.F.; Resources, U.S., C.F., H.A. and G.R.B.; Software, J.K. and J.R.; Supervision, C.F.; Validation, R.B., U.S., C.F., H.A. and G.R.B.; Visualization, J.K., J.S., J.R., R.B. and C.F.; Writing—original draft, J.K. and J.S.; Writing—review and editing, F.T., T.W.-R., J.R., R.B., U.S., C.F., H.A. and G.R.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study complies with the ethical standards of the Declaration of Helsinki and the institutional research committee and has been approved of by the ethics committee of the Charité—Universitätsmedizin Berlin (ID number: EA1/182/20; 30.09.2020).

**Informed Consent Statement:** Due to the retrospective character of this study, no informed consent was necessary.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy and ethical limitations.

**Conflicts of Interest:** The authors declare that there is no conflict of interest.

## References

1. Sai, K.K.S.; Zachar, Z.; Bingham, P.M.; Mintz, A. Metabolic PET Imaging in Oncology. *Am. J. Roentgenol.* **2017**, *209*, 270–276. [[CrossRef](#)] [[PubMed](#)]
2. Sugawara, Y.; Gutowski, T.D.; Fisher, S.J.; Brown, R.S.; Wahl, R.L. Uptake of positron emission tomography tracers in experimental bacterial infections: A comparative biodistribution study of radiolabeled FDG, thymidine, L-methionine,  $^{67}\text{Ga}$ -citrate, and  $^{125}\text{I}$ -HSA. *Eur. J. Nucl. Med.* **1999**, *26*, 333–341. [[CrossRef](#)] [[PubMed](#)]
3. Kung, B.T.; Seraj, S.M.; Zadeh, M.Z.; Rojulpote, C.; Kotheekar, E.; Ayubcha, C.; Ng, K.S.; Ng, K.K.; Au-Yong, T.K.; Werner, T.J.; et al. An update on the role of  $^{18}\text{F}$ -FDG-PET/CT in major infectious and inflammatory diseases. *Am. J. Nucl. Med. Mol. Imaging* **2019**, *9*, 255–273. [[PubMed](#)]
4. Blockmans, D.; Maes, A.; Stroobants, S.; Nuyts, J.; Bormans, G.; Knockaert, D.; Bobbaers, H.; Mortelmans, L. New arguments for a vasculitic nature of polymyalgia rheumatica using positron emission tomography. *Rheumatology* **1999**, *38*, 444–447. [[CrossRef](#)] [[PubMed](#)]
5. Tatò, F.; Hoffmann, U. Giant cell arteritis: A systemic vascular disease. *Vasc. Med.* **2008**, *13*, 127–140. [[CrossRef](#)] [[PubMed](#)]
6. Brack, A.; Martinez-Taboada, V.; Stanson, A.; Goronzy, J.J.; Weyand, C.M. Disease pattern in cranial and large-vessel giant cell arteritis. *Arthritis Rheum.* **1999**, *42*, 311–317. [[CrossRef](#)] [[PubMed](#)]
7. Barraclough, K.; Mallen, C.D.; Helliwell, T.; Hider, S.L.; Dasgupta, B. Diagnosis and management of giant cell arteritis. *Br. J. Gen. Pract.* **2012**, *62*, 329–330. [[CrossRef](#)]
8. Soussan, M.; Nicolas, P.; Schramm, C.; Katsahian, S.; Pop, G.; Fain, O.; Mekinian, A. Management of Large-Vessel Vasculitis With FDG-PET: A Systematic Literature Review and Meta-Analysis. *Medicine* **2015**, *94*, e622. [[CrossRef](#)]
9. Besson, F.L.; Parienti, J.-J.; Bienvenu, B.; Prior, J.O.; Costo, S.; Bouvard, G.; Agostini, D. Diagnostic performance of  $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography in giant cell arteritis: A systematic review and meta-analysis. *Eur. J. Nucl. Med.* **2011**, *38*, 1764–1772. [[CrossRef](#)]

10. Slart, R.H.J.A.; Writing Group. 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. [[CrossRef](#)]
11. Meignan, M.; Gallamini, A.; Haioun, C. Report on the First International Workshop on interim-PET scan in lymphoma. *Leuk. Lymphoma* **2009**, *50*, 1257–1260. [[CrossRef](#)]
12. Biggi, A.; Gallamini, A.; Chauvie, S.; Hutchings, M.; Kostakoglu, L.; Gregianin, M.; Meignan, M.; Malkowski, B.; Hofman, M.S.; Barrington, S.F. International Validation Study for Interim PET in ABVD-Treated, Advanced-Stage Hodgkin Lymphoma: Interpretation Criteria and Concordance Rate Among Reviewers. *J. Nucl. Med.* **2013**, *54*, 683–690. [[CrossRef](#)] [[PubMed](#)]
13. Furth, C.; Amthauer, H.; Hautzel, H.; Steffen, I.; Ruf, J.; Schiefer, J.; Schönberger, S.; Henze, G.; Grandt, R.; Hundsdoerfer, P.; et al. Evaluation of interim PET response criteria in paediatric Hodgkin's lymphoma—Results for dedicated assessment criteria in a blinded dual-centre read. *Ann. Oncol.* **2011**, *22*, 1198–1203. [[CrossRef](#)]
14. Dejaco, C.; Duftner, C.; Buttgereit, F.; Matteson, E.L.; Dasgupta, B. The spectrum of giant cell arteritis and polymyalgia rheumatica: Revisiting the concept of the disease. *Rheumatology* **2017**, *56*, 506–515. [[CrossRef](#)] [[PubMed](#)]
15. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33*, 159–174. [[CrossRef](#)] [[PubMed](#)]
16. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Med.* **2012**, *22*, 276–282. [[CrossRef](#)]
17. Watts, R.A.; Hatemi, G.; Burns, J.C.; Mohammad, A.J. Global epidemiology of vasculitis. *Nat. Rev. Rheumatol.* **2022**, *18*, 22–34. [[CrossRef](#)]
18. Prior, J.A.; Ranjbar, H.; Belcher, J.; Mackie, S.L.; Helliwell, T.; Liddle, J.; Mallen, C.D. Diagnostic delay for giant cell arteritis—A systematic review and meta-analysis. *BMC Med.* **2017**, *15*, 120. [[CrossRef](#)]
19. Ponte, C.; Martins-Martinho, J.; Luqmani, R.A. Diagnosis of giant cell arteritis. *Rheumatology* **2020**, *59*, iii5–iii16. [[CrossRef](#)]
20. Koster, M.J.; Matteson, E.L.; Warrington, K.J. Large-vessel giant cell arteritis: Diagnosis, monitoring and management. *Rheumatology* **2018**, *57*, ii32–ii42. [[CrossRef](#)]
21. Pugh, D.; Karabayas, M.; Basu, N.; Cid, M.C.; Goel, R.; Goodyear, C.S.; Grayson, P.C.; McAdoo, S.P.; Mason, J.C.; Owen, C.; et al. Large-vessel vasculitis. *Nat. Rev. Dis. Prim.* **2022**, *7*, 93. [[CrossRef](#)] [[PubMed](#)]
22. Kluge, R.; Chavdarova, L.; Hoffmann, M.; Kobe, C.; Malkowski, B.; Montravers, F.; Kurch, L.; Georgi, T.; Dietlein, M.; Wallace, W.H.; et al. Inter-Reader Reliability of Early FDG-PET/CT Response Assessment Using the Deauville Scale after 2 Cycles of Intensive Chemotherapy (OEPA) in Hodgkin's Lymphoma. *PLoS ONE* **2016**, *11*, e0149072. [[CrossRef](#)] [[PubMed](#)]
23. Furth, C.; Erdrich, A.S.; Steffen, I.G.; Stiebler, R.M.; Kahraman, D.; Kobe, C.; Schonberger, S.; Grandt, R.; Hundsdoerfer, P.; Hauptmann, K.; et al. Interim PET response criteria in paediatric non-Hodgkin's lymphoma. *Nuklearmedizin* **2013**, *52*, 148–156. [[CrossRef](#)]
24. Kobe, C.; Goergen, H.; Baues, C.; Kuhnert, G.; Voltin, C.-A.; Zijlstra, J.; Hoekstra, O.; Mettler, J.; Drzezga, A.; Engert, A.; et al. Outcome-based interpretation of early interim PET in advanced-stage Hodgkin lymphoma. *Blood* **2018**, *132*, 2273–2279. [[CrossRef](#)] [[PubMed](#)]
25. Johnson, P.; Federico, M.; Kirkwood, A.; Fossà, A.; Berkahn, L.; Carella, A.; D'Amore, F.; Enblad, G.; Franceschetto, A.; Fulham, M.; et al. Adapted Treatment Guided by Interim PET-CT Scan in Advanced Hodgkin's Lymphoma. *N. Engl. J. Med.* **2016**, *374*, 2419–2429. [[CrossRef](#)] [[PubMed](#)]
26. Radford, J.; Illidge, T.; Counsell, N.; Hancock, B.; Pettengell, R.; Johnson, P.; Wimperis, J.; Culligan, D.; Popova, B.; Smith, P.; et al. Results of a Trial of PET-Directed Therapy for Early-Stage Hodgkin's Lymphoma. *N. Engl. J. Med.* **2015**, *372*, 1598–1607. [[CrossRef](#)]
27. Cheson, B.D.; Fisher, R.I.; Barrington, S.F.; Cavalli, F.; Schwartz, L.H.; Zucca, E.; Lister, T.A. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: The Lugano classification. *J. Clin. Oncol.* **2014**, *32*, 3059–3068. [[CrossRef](#)]
28. Lee, Y.; Choi, S.; Ji, J.; Song, G. Diagnostic accuracy of <sup>18</sup>F-FDG PET or PET/CT for large vessel vasculitis: A meta-analysis. *Z. Für Rheumatol.* **2016**, *75*, 924–931. [[CrossRef](#)]
29. Besson, F.L.; de Boysson, H.; Parienti, J.-J.; Bouvard, G.; Biennu, B.; Agostini, D. Towards an optimal semiquantitative approach in giant cell arteritis: An <sup>18</sup>F-FDG PET/CT case-control study. *Eur. J. Nucl. Med. Mol. Imaging* **2014**, *41*, 155–166. [[CrossRef](#)]
30. Hautzel, H.; Sander, O.; Heinzel, A.; Schneider, M.; Müller, H.-W. Assessment of Large-Vessel Involvement in Giant Cell Arteritis with <sup>18</sup>F-FDG PET: Introducing an ROC-Analysis-Based Cutoff Ratio. *J. Nucl. Med.* **2008**, *49*, 1107–1113. [[CrossRef](#)]
31. Walter, M.A.; Melzer, R.A.; Schindler, C.; Müller-Brand, J.; Tyndall, A.; Nitzsche, E.U. The value of [<sup>18</sup>F]FDG-PET in the diagnosis of large-vessel vasculitis and the assessment of activity and extent of disease. *Eur. J. Nucl. Med.* **2005**, *32*, 674–681. [[CrossRef](#)] [[PubMed](#)]
32. Fuchs, M.; Briel, M.; Daikeler, T.; Walker, U.A.; Rasch, H.; Berg, S.; Ng, Q.K.T.; Raatz, H.; Jayne, D.; Kötter, I.; et al. The impact of <sup>18</sup>F-FDG PET on the management of patients with suspected large vessel vasculitis. *Eur. J. Nucl. Med.* **2012**, *39*, 344–353. [[CrossRef](#)] [[PubMed](#)]
33. Koenigkam-Santos, M.; Sharma, P.; Kalb, B.; Oshinski, J.N.; Weyand, C.M.; Goronzy, J.J.; Martin, D.R. Magnetic Resonance Angiography in Extracranial Giant Cell Arteritis. *JCR J. Clin. Rheumatol.* **2011**, *17*, 306–310. [[CrossRef](#)] [[PubMed](#)]
34. Meller, J.; Strutz, F.; Siefker, U.; Scheel, A.; Sahlmann, C.O.; Lehmann, K.; Conrad, M.; Vosschenrich, R. Early diagnosis and follow-up of aortitis with [<sup>18</sup>F]FDG PET and MRI. *Eur. J. Nucl. Med.* **2003**, *30*, 730–736. [[CrossRef](#)] [[PubMed](#)]

35. Lensen, K.D.F.; Comans, E.F.I.; Voskuyl, A.E.; van der Laken, C.J.; Brouwer, E.; Zwijnenburg, A.T.; Arias-Bouda, L.M.P.; Glaudemans, A.W.J.M.; Slart, R.H.J.A.; Smulders, Y.M. Large-Vessel Vasculitis: Interobserver Agreement and Diagnostic Accuracy of  $^{18}\text{F}$ -FDG-PET/CT. *BioMed Res. Int.* **2015**, *2015*, 914692. [[CrossRef](#)]
36. van der Geest, K.S.M.; Treglia, G.; Glaudemans, A.W.J.M.; Brouwer, E.; Sandovici, M.; Jamar, F.; Gheysens, O.; Slart, R.H.J.A. Diagnostic value of  $^{18}\text{F}$ FDG-PET/CT for treatment monitoring in large vessel vasculitis: A systematic review and meta-analysis. *Eur. J. Nucl. Med.* **2021**, *48*, 3886–3902. [[CrossRef](#)]
37. Mach, E. Über die Wirkung der räumlichen Verteilung des Lichtreizes auf die Netzhaut (I). In *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften; Mathematisch-Naturwissenschaftliche Classe 52*, K. k. Hof- und Staatsdruckerei: Vienna, Austria, 1865; pp. 303–322.
38. Boellaard, R.; Kobe, C.; Zijlstra, J.M.; Mikhaeel, N.G.; Johnson, P.W.; Müller, S.; Dührsen, U.; Hoekstra, O.S.; Barrington, S. Does PET Reconstruction Method Affect Deauville Scoring in Lymphoma Patients? *J. Nucl. Med.* **2018**, *59*, 1167–1169. [[CrossRef](#)]
39. Kuhnert, G.; Boellaard, R.; Sterzer, S.; Kahraman, D.; Scheffler, M.; Wolf, J.; Dietlein, M.; Drzezga, A.; Kobe, C. Impact of PET/CT image reconstruction methods and liver uptake normalization strategies on quantitative image analysis. *Eur. J. Nucl. Med. Mol. Imaging* **2016**, *43*, 249–258. [[CrossRef](#)]
40. Bucerius, J.; Mani, V.; Moncrieff, C.; Machac, J.; Fuster, V.; Farkouh, M.E.; Tawakol, A.; Rudd, J.H.F.; Fayad, Z.A. Optimizing  $^{18}\text{F}$ -FDG PET/CT imaging of vessel wall inflammation: The impact of  $^{18}\text{F}$ -FDG circulation time, injected dose, uptake parameters, and fasting blood glucose levels. *Eur. J. Nucl. Med. Mol. Imaging* **2013**, *41*, 369–383. [[CrossRef](#)]
41. Blomberg, B.A.; Thomassen, A.; Takx, R.A.P.; Hildebrandt, M.G.; Simonsen, J.A.; Buch-Olsen, K.M.; Diederichsen, A.C.P.; Mickley, H.; Alavi, A.; Høilund-Carlson, P.F. Delayed  $^{18}\text{F}$ -fluorodeoxyglucose PET/CT imaging improves quantitation of atherosclerotic plaque inflammation: Results from the CAMONA study. *J. Nucl. Cardiol.* **2014**, *21*, 588–597. [[CrossRef](#)]
42. Dejaco, C.; Ramiro, S.; Duftner, C.; Besson, F.L.; Bley, T.A.; Blockmans, D.; Brouwer, E.; Cimmino, M.A.; Clark, E.; Dasgupta, B.; et al. EULAR recommendations for the use of imaging in large vessel vasculitis in clinical practice. *Ann. Rheum. Dis.* **2018**, *77*, 636–643. [[CrossRef](#)] [[PubMed](#)]
43. Hunder, G.G.; Bloch, D.A.; Michel, B.A.; Stevens, M.B.; Arend, W.P.; Do, L.H.C.; Edworthy, S.M.; Fauci, A.S.; Leavitt, R.Y.; Lie, J.T.; et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. *Arthritis Rheum.* **1990**, *33*, 1122–1128. [[CrossRef](#)] [[PubMed](#)]
44. Stone, J.H.; Tuckwell, K.; Dimonaco, S.; Klearman, M.; Aringer, M.; Blockmans, D.; Brouwer, E.; Cid, M.C.; Dasgupta, B.; Rech, J.; et al. Trial of Tocilizumab in Giant-Cell Arteritis. *N. Engl. J. Med.* **2017**, *377*, 317–328. [[CrossRef](#)] [[PubMed](#)]
45. Huang, H.L.; Ngam, P.I.; Tan, K.M.; Ng, D.C.E.; Lim, S.T.; Chan, J.Y. The exact Deauville score, NABS score and high SUVmax predicts outcome in extranodal natural killer/T-cell lymphoma. *Ann. Nucl. Med.* **2021**, *35*, 557–568. [[CrossRef](#)]
46. Barrington, S.F.; Mikhaeel, N.G.; Kostakoglu, L.; Meignan, M.; Hutchings, M.; Müeller, S.P.; Schwartz, L.H.; Zucca, E.; Fisher, R.I.; Trotman, J.; et al. Role of Imaging in the Staging and Response Assessment of Lymphoma: Consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J. Clin. Oncol.* **2014**, *32*, 3048–3058. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.