^{68}Ga-DOTATATE PET/CT interobserver agreement for neuroendocrine tumor assessments: results from a prospective study on 50 patients

Wolfgang Peter Fendler[1], Martin Barrio[1], Claudio Spick[1], Martin Allen-Auerbach[1], Valentina Ambrosini[2], Matthias Benz[3], Christina Bluemel[4], Ravinder Kaur Grewal[5], Constantin Lapa[4], Matthias Miederer[6], Guillaume Nicolas[3], Tibor Schuster[7], Johannes Czernin[1], Ken Herrmann[1]

[1] Department of Molecular and Medical Pharmacology, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

[2] Nuclear Medicine, Department of Experimental Diagnostic and Specialized Medicine, University of Bologna and S. Orsola-Malpighi Hospital, Bologna, Italy

[3] Clinic of Radiology and Nuclear Medicine, University of Basel Hospital, Basel, Switzerland

[4] Department of Nuclear Medicine, Julius-Maximilians-University of Würzburg, Würzburg, Germany

[5] Molecular Imaging and Therapy Service, Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[6] Department of Nuclear Medicine, University Medical Center Mainz, Mainz, Germany

[7] Department of Family Medicine, McGill University, Montreal, Quebec, Canada

First author:

Wolfgang Peter Fendler, M.D.

University of California at Los Angeles

Ahmanson Translational Imaging Division

10833 Le Conte Ave.

200 Medical Plaza

Ste. B114-61

Los Angeles

CA 90095-7370

E-mail: wfendler@mednet.ucla.edu

Corresponding author:

Ken Herrmann, M.D.

University of California at Los Angeles

Ahmanson Translational Imaging Division

10833 Le Conte Ave.

200 Medical Plaza

Ste. B114-61

Los Angeles

CA 90095-7370

E-mail: kherrmann@mednet.ucla.edu

ABSTRACT

We evaluated the observer agreement for ^{68}Ga-DOTATATE PET/CT study interpretations in patients with neuroendocrine tumors (NET). **Methods:** ^{68}Ga-DOTATATE PET/CT was performed in 50 patients with known or suspected NET of the small bowel (n = 19), pancreas (n = 14), lung (n = 4) or other location (n = 13). Images were reviewed by seven observers who used a standardized approach for image interpretation. Observers were classified as having low (<500 scans or <5 years experience with ^{68}Ga-DOTATATE PET/CT; n = 4) or high level of experience (≥500 scans and ≥5 years experience with ^{68}Ga-DOTATATE PET/CT; n = 3). Interpretation by the primary nuclear medicine physician un-blinded to all clinical and imaging data served as reference standard. Interobserver agreement was determined by Cohen's κ and intraclass correlation coefficient (ICC) with corresponding 95% confidence interval (CI). **Results:** Interobserver agreement was substantial and the median number of false findings (FF) was low for the overall scan result; i.e. positive versus negative study (κ = 0.80, 95%CI 0.74-0.86; FF = 3), organ involvement (κ = 0.70, 95%CI 0.64-0.76; FF = 5), and lymph node involvement (κ = 0.71, 95%CI 0.65-0.78; FF = 6). The interobserver agreement was substantial to almost-perfect and the average absolute difference (Δ) to the reference reader was low for number of organ and lymph node metastases (ICC = 0.84, 95%CI 0.77-0.89, Δ = 0.45 and ICC = 0.77, 95%CI 0.69-0.84, Δ = 0.45), tumor SUV_{max} (ICC = 0.99, 95%CI 0.97-0.99; Δ = 0.44) and reference SUV (SUV_{mean} spleen: ICC = 0.81, Δ = 1.10; SUV_{max} liver ICC = 0.79, Δ = 0.62). Interpretations of the appropriateness for peptide-receptor radionuclide therapy (PRRT) varied more significantly among observers (κ = 0.64, 95%CI 0.57-0.70) and a higher frequency of false positive recommendations for PRRT occurred in observers with low versus high levels of experience (range, 7-12 versus 4-8). **Conclusion:** The interpretation of ^{68}Ga-DOTATATE PET/CT for NET staging is consistent among readers with low and high levels of experience. However, image based recommendations for or against PRRT require experience and training.

INTRODUCTION

Overexpression of cell surface somatotstatin receptors (SSRs) in well-differentiated neuroendocrine tumors (NETs) can be exploited for imaging and therapy with radiolabeled somatostatin analogues. SSR scintigraphy using [111]Indium-Octreotide (OcetreoScan®) has been available for more than 20 years. Over the past decades scintigraphy was gradually replaced by ^{68}Ga-DOTATOC or ^{68}Ga-DOTATATE Positron Emission Tomography/Computed Tomography (PET/CT) imaging because of their superior accuracy for NET staging (*1-4*). PET/CT using ^{68}Ga-labeled somatostatin analogues is now considered the gold standard and its superiority over scintigraphy is also emphasized in the European Neuroendocrine Tumor Society guidelines (*5*).

^{68}Ga-DOTATATE PET/CT demonstrates high accuracy for NET staging and is an important companion diagnostic to the highly effective ^{177}Lu-DOTATATE PRRT (*6*). High remission rates after PRRT were positively correlated with intense [111]Indium-Octreotide uptake on pre-therapy SSR imaging using a liver-based 4-point scale for tracer accumulation (*7*). ^{68}Ga-DOTATATE uptake might thus similar to the Krenning scale predict likelihood of response to PRRT. ^{68}Ga-DOTATATE has now received approval by the US Food and Drug Administration for NET imaging. However, little is known about the inter-reader differences for ^{68}Ga-DOTATATE PET/CT interpretations. The overall value of an imaging method is associated with the degree of reader agreement. Knowledge of interobserver variability and reproducibility is therefore essential for interpreting study results and design of future trials. The aim of this study was to determine interobserver agreement for interpretations of ^{68}Ga-DOTATATE PET/CT and to compare findings among readers with low and high levels of experience.

PATIENTS AND METHODS

Patients and Image Acquisition

From June 2013 until March 2014, 50 patients with known or suspected NET of the small bowel (n = 19), pancreas (n = 14), lung (n = 4) or other location (n = 13) were prospectively

recruited under a Food and Drug Administration approved Investigational New Drug application, a subgroup of our previous publication. Patients were referred for NET staging (n = 10) or re-staging (n = 40). All patients were included in a previously published study on the impact of ^{68}Ga-DOTATATE PET/CT on patient management (8). The prospective study was approved by the University of California, Los Angeles Institutional Review Board (IRB) and all subjects signed a written informed consent. Patient preparation and image acquisition were performed as previously described (8). In brief, ^{68}Ga-DOTATATE was injected intravenously at a dose of 190 ± 17 MBq (5.1 ± 0.5 mCi; range, 130-211 MBq [3.5-5.7 mCi]). Tracer uptake period of 60 min (mean, 62 ± 7 min) was allowed before imaging with a Biograph True Point 64 or Biograph mCT device (Siemens). For anatomic correlation and identification of organ lesions, intravenous contrast was administered in 47 of 50 patients (94%) using 90 to 115 mL of Omnipaque 350 (GE Healthcare) at a flow rate of 2 mL/s. Oral contrast was given in all patients using approximately 600 ml of Barium Sulfate (Readi-Cat 2, Bracco) within one hour before the scan.

Observer

Anonymized PET/CT images of 50 patients included in the study (one scan per patient) were electronically submitted to 7 nuclear medicine physicians from 5 centers located in Europe (n = 4) or North America (n = 1). Data included standard DICOM files of CT, attenuation-corrected and uncorrected PET images. All physicians had at least five years of prior experience with interpreting oncologic ^{18}F-FDG PET/CT scans. All centers had been performing ^{68}Ga-DOTATATE PET/CT on a regular basis for at least two years and all physicians had prior experience with ^{68}Ga-DOTATATE PET/CT. Observers reported years of experience and the total number of ^{68}Ga-DOTATATE PET/CT scans that they had interpreted before. Based on these reports observers were grouped into the following experience categories: (a) low level: less than 500 scans or less than 5 years experience with ^{68}Ga-DOTATATE PET/CT (n = 4); (b) high level: at least 500 scans and more than 5 years experience with ^{68}Ga-DOTATATE PET/CT (n = 3).

Visual Interpretation

A written guide for image interpretation and two patient examples were provided to each observer (Supplemental Material). The following patient information was disclosed to each observer before image interpretation: Indication (staging/re-staging), gender (male/female), age (years), height (cm), weight (kg), injected dose (mCi), uptake time (min), CT protocol (contrast-enhanced/non-enhanced), prior therapy (yes/no), prior chemotherapy (yes/no), prior PRRT (yes/no), prior surgery (yes/no), and prior octreotide therapy (yes/no). Observers were blinded to all other clinical data.

Visual image interpretation was reported for the following pre-defined categories: Overall scan result for presence or absence of disease, SSR density in NET tissue (none=0/low=1/intermediate=2/high=3), indication for PRRT (yes/no), organs affected (yes/no), number of organs affected (0,1,2,3,4,≥5), number of organ metastases detected (0,1,2,3,4,≥5), lymph nodes affected (yes/no), number of lymph node regions affected (0,1,2,3,4,≥5), and number of lymph node metastases detected (0,1,2,3,4,≥5). Criteria for PRRT indication were among others intense tracer uptake of tumor lesions and metastatic spread.

SUV Measurement

Each observer recorded series number, image number, location and SUV_{max} for up to three target lesions. Lesions were chosen from each diseased organ system following pre-defined criteria as outlined in the written guide for image interpretation (Supplemental Material).

Each observer measured maximum standardized uptake value (SUV_{max}) and average SUV (SUV_{mean}) using a five centimeter diameter circular region of interest (ROI) placed in a lesion free region of the right hepatic lobe and a two centimeter diameter circular ROI in lesion

free splenic parenchyma. To ascertain consistency readers were provided with a PET/CT dataset and SUV data for one test patient for comparison.

Reference Standard and statistical Analyses

Data acquisition and analyses were performed prospectively. Scan interpretations by a University of California, Los Angeles physician (MAA) who was unblinded for all baseline and follow-up clinical information served as reference standard. For binary data, agreement among each observer and the reference reader was evaluated using Cohen's κ (*9*). Overall agreement using pooled observer data was evaluated using generalized estimation equations (*10*). For non-binary data, agreement among all observers was evaluated by intraclass correlation coefficient using two-way mixed model for absolute agreement (single measures) (*11*). To calculate intraclass correlation coefficient for tumor SUV_{max} one target lesion, i.e. the lesion reported by most observers, was chosen per patient. Ninety-five percent confidence intervals (CIs) are reported for κ and intraclass correlation coefficient values. Interpretation of κ and intraclass correlation coefficient was based on a classification provided by Landis and Koch (*12*): 0.0, poor; 0.0–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; 0.81–1.00, almost-perfect reproducibility.

Discrepancies in quantitative ratings among observers were expressed as mean difference (Δ) ± standard deviation (SD). Statistical analyses were performed using R software (R Core Team 2015, R Foundation for Statistical Computing, Vienna, Austria) with the package "irr" (Gamer et al, version 0.84) for generalized estimation equation modelling and SPSS (version 15.0, SPSS Inc., Chicago, Illinois, USA) for all other statistical analyses.

RESULTS

Patient Characteristics

Table 1 summarizes patient characteristics. 68Ga-DOTATATE PET/CT studies were rated positive for NET by the reference reader in 37 of 50 (74%) patients. Twenty-one patients (42%) had N1, while 34 (68%) were staged as M1 by the reference reader.

Visual Interpretation

The interobserver agreement for visual interpretation is shown in Table 2. Reproducibility was substantial to almost-perfect for overall scan result, organ involvement, lymph node involvement and respective subitems, i.e. number of organs/organ metastases/LN areas/LN metastases (each intraclass correlation coefficient/κ ≥0.70). Mean absolute delta to the reference reader was low for number of organs, lymph node areas or metastases (each Δ<0.5) and not relevantly different among observers with low versus high level of experience. However, the interobserver agreement for PRRT indication ranged from only moderate to substantial (κ = 0.64, 95%CI 0.57-0.70).

False positive and negative findings as well as level of agreement between individual observers and the reference reader are listed in Table 3. For the overall scan result, organ and lymph node involvement observers from either level of experience demonstrated a low frequency of false positive or negative findings (range 0 to 6).

A patient example for false positive overall scan result by five of seven observers is given in Figure 1. Three of four observers with low level of experience had a fair agreement with the reference reader for PRRT indication. The number of erroneous recommendations for PRRT was higher for observers with low (range 7 to 12) than those with high levels of experience (range 4 to 8). Details on individual test performance as well as sensitivity and specificity values are available in Supplemental Table 1.

SUV Measurements

Interobserver agreement for SUV measurements is given in Table 4. Agreement was almost-perfect for tumor SUV_{max} (intraclass correlation coefficient 0.99). SUV_{max} liver and SUV_{mean} spleen were highly reproducible (intraclass correlation coefficient 0.79 and 0.81 respectively) with a low mean absolute differences (each $\Delta<1.2$) when compared to SUV measurements of the reference reader. Mean absolute delta was comparable between observers with low versus high levels of experience. Figure 2 illustrates agreement among individual SUV measurements.

DISCUSSION

We analyzed the reproducibility of ^{68}Ga-DOTATATE PET/CT scan interpretations in 50 patients by comparing assessments from one non-blinded in-house reference reader with those of seven external observers. All physicians had extensive experience in the interpretation of oncological ^{18}F-FDG PET/CT. Based on their experience with reading ^{68}Ga-DOTATATE PET/CT scans physicians were grouped into having low or high levels of experience. The overall scan result and most of the individual parameters, including SUV analysis were highly reproducible. Disagreements occurred at times in drawing conclusions from the image findings: Observers with lower levels of experience arrived at times at different PRRT recommendations than the more experienced observers and the reference reader.

^{68}Ga-DOTATATE PET/CT scan interpretation is not without pitfalls. Up to 70% of high-grade NET lesions are ^{68}Ga-DOTATATE PET-negative due to low or even absent SSR expression (*13*). On the other hand inflammation with recruitment of SSR expressing macrophages may lead to false positive findings (*14*). Physiologic uptake in adrenal glands, the pituitary gland and the uncinate process of the pancreas potentially mask NET lesions or can be misinterpreted as tumor tissue (*15*). Other pitfalls include uptake at sites of osteoblastic activity, splenules, splenosis, or SSR expression in hemangioma or other benign and malignant tumors

of non-neuroendocrine origin (*16,17*). All these processes may result in interpretative errors that ultimately limit the accuracy of ^{68}Ga-DOTATATE PET/CT. To reduce the frequency of errors most study protocols include consensus image analysis by multiple readers. However, the level of interobserver agreement cannot be discerned from consensus readings (*8,13,18-20*). Here we provide evidence for substantial to almost-perfect agreement among seven readers with varying experience levels. Agreement was determined for both visual and semi-quantitative analyses in multiple, pre-defined categories. Previous reports included reproducibility as secondary or tertiary endpoints and compared findings from no more than three readers with comparable or unknown level of experience. Deppen et al reports almost-perfect reproducibility (κ = 0.82) between two blinded and one non-blinded reader of ^{68}Ga-DOTATATE PET/CT in 78 patients with pulmonary or gastro-entero-pancreatic NET (*3*). Ruf et al analyzed agreement between two readers after separate analysis of ^{68}Ga-DOTATOC PET and triple-phase CT in 51 NET patients. Agreement was substantial for PET (κ = 0.77) however only fair to moderate for triple-phase CT and single CT phases (*21*). Two trials employed separate analyses by independent readers, but did not report reproducibility (*22,23*).

Interobserver agreement is an important aspect of clinical applicability. Complex protocols can be associated with reduced inter-reader agreement as has been demonstrated for triple-phase CT (*21*) and single-/multi-sequence MRI (*24,25*) for NET staging. Here we demonstrate that both visual and semi-quantitative ^{68}Ga-DOTATATE PET/CT analysis is highly reproducible among observers with both low and high experience. Our findings indicate a high inter-reader reliability of ^{68}Ga-DOTATATE PET/CT for NET staging in a clinical and research setting, even if image interpretation is performed by less experienced readers. Interestingly observers with lower experience levels inappropriately recommended PRRT with a higher frequency than highly experienced readers. ^{68}Ga-DOTATATE PET/CT should thus be reviewed by experienced readers if PRRT is considered in NET patients.

The present study has several limitations. First, observers reviewed the ^{68}Ga-DOTATATE PET/CT and a limited set of patient information, whereas the reference reader had access to all available clinical and image data. Clinical data such as type and duration of prior therapy are important to determine appropriateness for PRRT. Lack of clinical data creates a disadvantage which may have led to a higher frequency of PRRT being inappropriately indicated by the observers. Second, observers were grouped based on experience with ^{68}Ga-DOTATATE PET/CT interpretation. However it should be noted that the skills of a reader are determined by multiple additional factors, including clinical knowledge and experience with other imaging modalities. Thus unknown factors might have led to an underestimation of the true expertise, especially for readers in the low experience group.

CONCLUSION

Both visual and semi-quantitative analysis of ^{68}Ga-DOTATATE PET/CT is highly reproducible among readers with varying experience. Diagnostic information gained from ^{68}Ga-DOTATATE PET/CT in an appropriate clinical or research setting can thus be considered reliable. ^{68}Ga-DOTATATE PET/CT interpretation should be performed by experienced readers if PRRT is considered.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING

REFERENCES

1.   Gabriel M, Decristoforo C, Kendler D, et al. 68Ga-DOTA-Tyr3-octreotide PET in neuroendocrine tumors: comparison with somatostatin receptor scintigraphy and CT. *J Nucl Med.* 2007;48:508-518.

2.   Srirajaskanthan R, Kayani I, Quigley AM, Soh J, Caplin ME, Bomanji J. The role of 68Ga-DOTATATE PET in patients with neuroendocrine tumors and negative or equivocal findings on 111In-DTPA-octreotide scintigraphy. *J Nucl Med.* 2010;51:875-882.

3.   Deppen SA, Liu E, Blume JD, et al. Safety and Efficacy of 68Ga-DOTATATE PET/CT for Diagnosis, Staging and Treatment Management of Neuroendocrine Tumors. *J Nucl Med.* 2016.

4.   Deppen SA, Blume J, Bobbey AJ, et al. 68Ga-DOTATATE Compared with 111In-DTPA-Octreotide and Conventional Imaging for Pulmonary and Gastroenteropancreatic Neuroendocrine Tumors: A Systematic Review and Meta-Analysis. *J Nucl Med.* 2016;57:872-878.

5.   Pape UF, Perren A, Niederle B, et al. ENETS Consensus Guidelines for the management of patients with neuroendocrine neoplasms from the jejuno-ileum and the appendix including goblet cell carcinomas. *Neuroendocrinology.* 2012;95:135-156.

6.   Ruszniewski P. 177Lu-DOTATATE significantly improves progression-free survival in patients with mid-gut neuroendocrine tumours: results of the phase III NETTER-1 trial. *ECC 2015.* 2015;abstract LBA 6.

7.   Kwekkeboom DJ, Teunissen JJ, Bakker WH, et al. Radiolabeled somatostatin analog [177Lu-DOTA0,Tyr3]octreotate in patients with endocrine gastroenteropancreatic tumors. *J Clin Oncol.* 2005;23:2754-2762.

8.   Herrmann K, Czernin J, Wolin EM, et al. Impact of 68Ga-DOTATATE PET/CT on the management of neuroendocrine tumors: the referring physician's perspective. *J Nucl Med.* 2015;56:70-75.

9.   Cohen JA. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:37-46.

10.   Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42:121-130.

11.   Scheffe H. The Analysis of Variance. New York, NY: Wiley. 1959:221-260.

12.   Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-174.

13.   Kayani I, Bomanji JB, Groves A, et al. Functional imaging of neuroendocrine tumors with combined PET/CT using 68Ga-DOTATATE (DOTA-DPhe1,Tyr3-octreotate) and 18F-FDG. *Cancer.* 2008;112:2447-2455.

14.   Armani C, Catalani E, Balbarini A, Bagnoli P, Cervia D. Expression, pharmacology, and functional role of somatostatin receptor subtypes 1 and 2 in human macrophages. *J Leukoc Biol.* 2007;81:845-855.

15.   Mapelli P, Tam HH, Sharma R, Aboagye EO, Al-Nahhas A. Frequency and significance of physiological versus pathological uptake of 68Ga-DOTATATE in the pancreas: validation with morphological imaging. *Nucl Med Commun.* 2014;35:613-619.

16.   Hofman MS, Lau WF, Hicks RJ. Somatostatin receptor imaging with 68Ga DOTATATE PET/CT: clinical utility, normal patterns, pearls, and pitfalls in interpretation. *Radiographics.* 2015;35:500-516.

17.   Brogsitter C, Hofmockel T, Kotzerke J. (68)Ga DOTATATE uptake in vertebral hemangioma. *Clin Nucl Med.* 2014;39:462-463.

18.   Naswa N, Sharma P, Kumar A, et al. Gallium-68-DOTA-NOC PET/CT of patients with gastroenteropancreatic neuroendocrine tumors: a prospective single-center study. *AJR Am J Roentgenol.* 2011;197:1221-1228.

19.   Haug AR, Cindea-Drimus R, Auernhammer CJ, et al. The role of 68Ga-DOTATATE PET/CT in suspected neuroendocrine tumors. *J Nucl Med.* 2012;53:1686-1692.

20.   Haug AR, Auernhammer CJ, Wangler B, et al. 68Ga-DOTATATE PET/CT for the early prediction of response to somatostatin receptor-mediated radionuclide therapy in patients with well-differentiated neuroendocrine tumors. *J Nucl Med.* 2010;51:1349-1356.

21.   Ruf J, Schiefer J, Furth C, et al. 68Ga-DOTATOC PET/CT of neuroendocrine tumors: spotlight on the CT phases of a triple-phase protocol. *J Nucl Med.* 2011;52:697-704.

22.   Sadowski SM, Neychev V, Millo C, et al. Prospective Study of 68Ga-DOTATATE Positron Emission Tomography/Computed Tomography for Detecting Gastro-Entero-Pancreatic Neuroendocrine Tumors and Unknown Primary Sites. *J Clin Oncol.* 2016;34:588-596.

23.   Buchmann I, Henze M, Engelbrecht S, et al. Comparison of 68Ga-DOTATOC PET and 111In-DTPAOC (Octreoscan) SPECT in patients with neuroendocrine tumours. *Eur J Nucl Med Mol Imaging.* 2007;34:1617-1626.

24.   Brenner R, Metens T, Bali M, Demetter P, Matos C. Pancreatic neuroendocrine tumor: added value of fusion of T2-weighted imaging and high b-value diffusion-weighted imaging for tumor detection. *Eur J Radiol.* 2012;81:e746-749.

25.   De Robertis R, D'Onofrio M, Zamboni G, et al. Pancreatic Neuroendocrine Neoplasms: Clinical Value of Diffusion-Weighted Imaging. *Neuroendocrinology.* 2015;103.
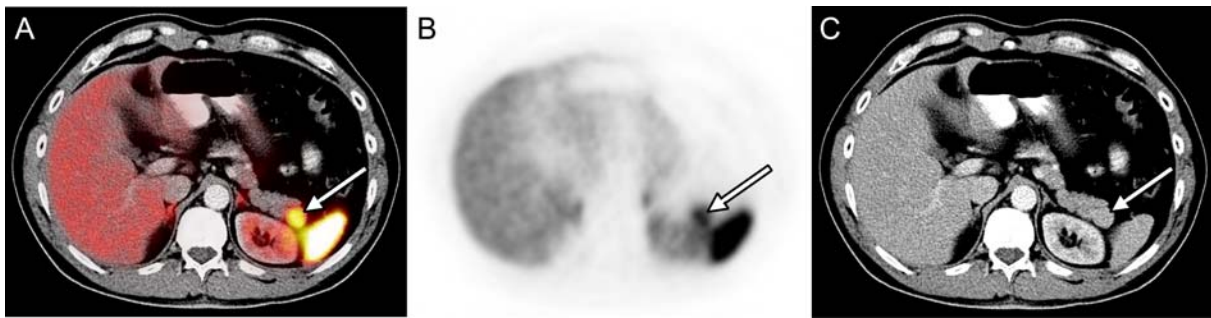
FIGURE LEGENDS

Figure 1. False positive finding for the overall scan result in a patient with repeated flushing. Axial fused 68Ga-DOTATATE PET/CT (A), PET (B) and CT (C) are shown. Focal 68Ga-DOTATATE uptake (arrow) in the region of the pancreatic tail was judged for NET by five of seven observers (false positive). Based on clinical data and MRI follow-up the reference reader ruled out NET and confirmed the presence of splenule abutting the tail of the pancreas.
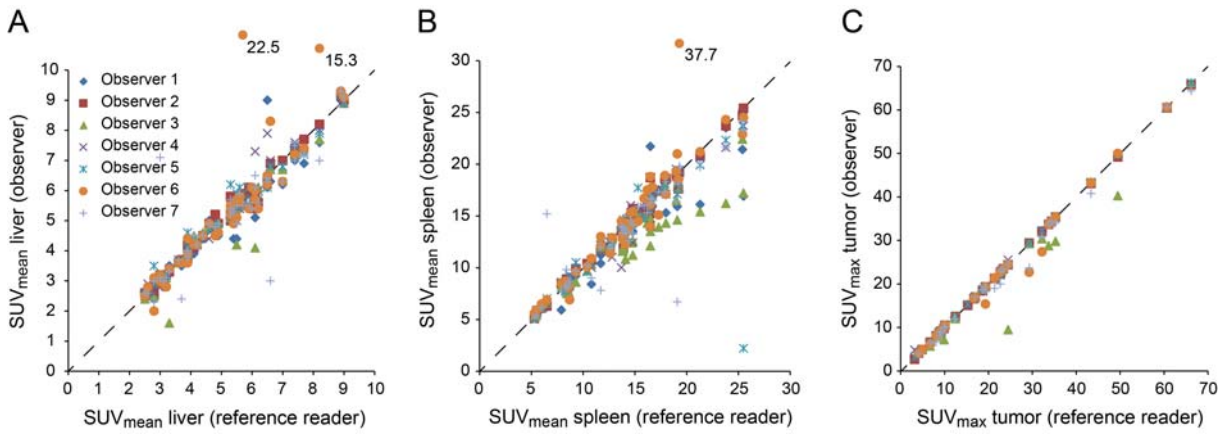
Figure 2. Interobserver agreement for SUV$_{mean}$ liver (A), SUV$_{mean}$ spleen (B) and SUV$_{max}$ tumor (C). SUV values were sorted by SUV obtained from the reference reader. Dashed diagonal lines indicate perfect agreement. Three Y-axis outliers were drawn in a relative position outside the scale and the absolute Y-value is given.

TABLES

Patient characteristics (n = 50)		Median (range) or absolute number (percent)
Age (years)		61 (27-78)
Sex		
	Female	25 (50%)
	Male	25 (50%)
Location of primary		
	Small bowel	19 (38%)
	Pancreas	14 (28%)
	Other	13 (26%)
	Lung	4 (8%)
Indication		
	Staging	10 (20%)
	Re-staging	40 (80%)
Scan positive for NET		37 (74%)
N1 stage		21 (42%)
M1 stage		34 (68%)
Prior therapy		
	CTx	11 (22%)
	PRRT	3 (6%)
	Surgery	30 (60%)
	Octreotide	22 (44%)

Table 1. Patient characteristics. Positive ⁶⁸Ga-DOTATATE PET/CT and N/M stage was determined by the reference reader. Abbreviations: CTx, chemotherapy; PRRT, peptide receptor radionuclide therapy.

	Agreement		95% CI	Mean Δ±SD, low group	Mean Δ±SD, high group
Overall scan result (pos/neg)	κ:	0.80	0.74-0.86		
PRRT indication (pos/neg)	κ:	0.64	0.57-0.70		
SSR density (0-3)	ICC:	0.84	0.77-0.89	0.30±0.64	0.31±0.65
Organ involvement (pos/neg)	κ:	0.70	0.64-0.76		
Number of organs (0-4, ≥5)	ICC:	0.75	0.66-0.83	0.42±0.67	0.39±0.66
Number of organ met. (0-4, ≥5)	ICC:	0.84	0.77-0.89	0.42±0.95	0.47±1.09
LN involvement	κ:	0.71	0.65-0.78		
Number of LN areas (0-4, ≥5)	ICC:	0.79	0.71-0.86	0.23±0.46	0.28±0.63
Number of LN met. (0-4, ≥5)	ICC:	0.77	0.69-0.84	0.42±0.90	0.47±1.00

Table 2. Interobserver agreement for visual image interpretation. Cohen's κ and intraclass correlation coefficient (ICC) are given. Mean absolute difference (Δ) when compared to findings of the reference reader was calculated separately for readers with low versus high experience. Abbreviations: SD, standard deviation; Pos, positive; neg, negative; PRRT, peptide receptor radionuclide therapy; SSR, somatostatin receptor; met, metastases; LN, lymph node.

Obs	Exp	Overall scan result			PRRT indication			Organ involvement			LN involvement		
		FP	FN	κ	FP	FN	κ	FP	FN	κ	FP	FN	κ
1	Low	1	2	0.85	7	1	0.68	2	0	0.91	6	2	0.68
2	Low	3	1	0.78	11	4	0.41	3	3	0.72	3	3	0.75
3	Low	2	1	0.84	12	1	0.50	4	1	0.76	1	2	0.88
4	Low	1	0	0.95	10	1	0.57	2	1	0.86	2	2	0.84
5	High	0	0	1.00	4	2	0.76	4	0	0.80	6	1	0.72
6	High	2	3	0.75	8	1	0.65	4	5	0.59	2	2	0.84
7	High	1	5	0.72	4	2	0.76	2	3	0.77	4	3	0.72

Table 3. False findings for each observer and agreement with the reference reader for evaluation of n = 50 cases. Abbreviations: Obs, observer number; Exp, experience level; FP, false positive; FN, false negative; PRRT, peptide receptor radionuclide therapy; LN, lymph node.

	ICC	95% CI	Mean Δ±SD, low group	Mean Δ±SD, high group
SUV$_{mean}$ liver	0.69	0.58 - 0.79	0.24±0.34	0.47±1.62
SUV$_{max}$ liver	0.79	0.70 - 0.86	0.46±0.61	0.78±1.20
SUV$_{mean}$ spleen	0.81	0.73 - 0.88	1.04±1.47	1.17±2.81
SUV$_{max}$ spleen	0.65	0.54 - 0.76	1.10±4.70	1.07±2.25
SUV$_{max}$ tumor	0.99	0.97-0.99	0.42±1.73	0.46±1.20

Table 4. Interobserver agreement for SUV values. Intraclass correlation coefficient (ICC) is given. Mean absolute difference (Δ) when compared to findings of the reference reader was calculated separately for observer with low versus high experience.