

# <sup>68</sup>Ga-FAPI PET/CT Interobserver Agreement on Tumor Assessment: An International Multicenter Prospective Study

Riccardo Mei<sup>\*1,2</sup>, Lukas Kessler<sup>\*1,3</sup>, Kim M. Pabst<sup>1,3</sup>, Manuel Weber<sup>1,3</sup>, Christian Schimdkonz<sup>4</sup>, Christoph Rischpler<sup>1,3</sup>, Helle Damgaard Zacho<sup>5</sup>, Thomas Hope<sup>6</sup>, Sarah M. Schwarzenböck<sup>7</sup>, Martin Allen-Auerbach<sup>8,9</sup>, Louise Emmett<sup>10</sup>, Justin Ferdinandus<sup>11</sup>, Marcus Unterrainer<sup>12</sup>, Benedikt M. Schaarschmidt<sup>13</sup>, Lale Umutlu<sup>13</sup>, Andrea Farolfi<sup>2</sup>, Paolo Castellucci<sup>2</sup>, Cristina Nanni<sup>2</sup>, Silvi Telo<sup>2</sup>, Stefano Fanti<sup>2</sup>, Ken Herrmann<sup>1,3</sup>, and Wolfgang P. Fendler<sup>1,3</sup>

<sup>1</sup>Department of Nuclear Medicine, University Hospital Essen, University of Duisburg–Essen, Essen, Germany; <sup>2</sup>Nuclear Medicine Department, IRCCS Azienda Ospedaliero–Universitaria di Bologna, Bologna, Italy; <sup>3</sup>German Cancer Consortium, Partner Site University Hospital Essen, and German Cancer Research Center, Essen, Germany; <sup>4</sup>Department of Nuclear Medicine, University Hospital Erlangen, Erlangen, Germany; <sup>5</sup>Department of Clinical Medicine, Aalborg University Hospital, Aalborg, Denmark; <sup>6</sup>Department of Radiology and Biomedical Imaging, University of California, San Francisco, California; <sup>7</sup>Department of Nuclear Medicine, Rostock University Medical Centre, Rostock, Germany; <sup>8</sup>Ahmanson Translational Theranostics Division, Department of Molecular and Medical Pharmacology, David Geffen School of Medicine, UCLA, Los Angeles, California; <sup>9</sup>Institute of Urologic Oncology, David Geffen School of Medicine, UCLA, Los Angeles, California; <sup>10</sup>Department of Theranostics and Nuclear Medicine, St. Vincent's Hospital, and Faculty of Medicine, UNSW Sydney, Sydney, New South Wales, Australia; <sup>11</sup>Department of Internal Medicine I, University of Cologne, Cologne, Germany; <sup>12</sup>Department of Radiology, University Hospital LMU Munich, Munich, Germany; and <sup>13</sup>Department of Diagnostic and Interventional Radiology and Neuroradiology, University Hospital Essen, University of Duisburg–Essen, Essen, Germany

<sup>68</sup>Ga-fibroblast activation protein inhibitors (FAPIs) are promising radiotracers for cancer imaging, with emerging data in the recent years. Nonetheless, the interobserver agreement on <sup>68</sup>Ga-FAPI PET/CT study interpretations in cancer patients remains poorly understood. **Methods:** <sup>68</sup>Ga-FAPI PET/CT was performed on 50 patients with various tumor entities (sarcoma [ $n = 10$ ], colorectal cancer [ $n = 10$ ], pancreatic adenocarcinoma [ $n = 10$ ], genitourinary cancer [ $n = 10$ ], and other types of cancer [ $n = 10$ ]). Fifteen masked observers reviewed and interpreted the images using a standardized approach for local, local nodal, and metastatic involvement. Observers were grouped by experience as having a low (<30 prior <sup>68</sup>Ga-FAPI PET/CT studies;  $n = 5$ ), intermediate (30–300 studies;  $n = 5$ ), or high level of experience (>300 studies;  $n = 5$ ). Two independent readers with a high level of experience and unmasked to clinical information, histopathology, tumor markers, and follow-up imaging (CT/MRI or PET/CT) served as the standard of reference (SOR). Observer groups were compared by overall agreement (percentage of patients matching SOR) and Fleiss  $\kappa$  with mean and corresponding 95% CI. We defined acceptable agreement as a  $\kappa$  value of at least 0.6 (substantial or higher) and acceptable accuracy as at least 80%. **Results:** Highly experienced observers agreed substantially on all categories (primary tumor:  $\kappa = 0.71$ ; 95% CI, 0.71–0.71; local nodal involvement:  $\kappa = 0.62$ ; 95% CI, 0.61–0.62; distant metastasis:  $\kappa = 0.75$ ; 95% CI, 0.75–0.75), whereas observers with intermediate experience showed substantial agreement on primary tumor ( $\kappa = 0.73$ ; 95% CI, 0.73–0.73) and distant metastasis ( $\kappa = 0.65$ ; 95% CI, 0.65–0.65) but moderate agreement on local nodal stages ( $\kappa = 0.55$ ; 95% CI, 0.55–0.55). Observers with low experience had moderate agreement on all categories (primary tumor:  $\kappa = 0.57$ ; 95% CI, 0.57–0.58; local nodal involvement:  $\kappa = 0.51$ ; 95% CI, 0.51–0.52; distant metastasis:  $\kappa = 0.54$ ; 95% CI, 0.53–0.54). Compared with SOR, the accuracy for readers with

high, intermediate, and low experience was 85%, 83%, and 78%, respectively. In summary, only highly experienced readers showed substantial agreement and a diagnostic accuracy of at least 80% in all categories. **Conclusion:** The interpretation of <sup>68</sup>Ga-FAPI PET/CT for cancer imaging had substantial reproducibility and accuracy among highly experienced observers only, especially for local nodal and metastatic assessments. Therefore, for accurate interpretation of different tumor entities and pitfalls, we recommend training or experience with at least 300 representative scans for future clinical readers.

**Key Words:** cancer imaging; FAPI; fibroblast activation protein; PET; interobserver agreement

**J Nucl Med 2023; 00:1–6**

DOI: 10.2967/jnumed.122.265245

**F**ibroblast activation protein is expressed by carcinoma-associated fibroblasts and cells of certain solid tumors (*1*). Radiolabeled fibroblast activation protein inhibitors (FAPIs) such as <sup>68</sup>Ga-FAPI-46 have been developed as novel theranostic tools for cancer (*2–5*). Increasing data suggest a high sensitivity for tumor detection and less background uptake than with <sup>18</sup>F-FDG (*6–8*). Mainly retrospective trials have demonstrated superior detection rates and higher accuracy for the localization for various tumor entities, but larger cohorts and prospective data are lacking. Nonetheless, multiple prospective trials to evaluate the accuracy and management impact of <sup>68</sup>Ga-FAPI PET/CT are currently under way at our site and worldwide (NCT05160051, NCT04023240, NCT05262855) (*7,9,10*).

To serve as a reliable clinical and research tool in the future, findings from <sup>68</sup>Ga-FAPI PET/CT must be accurate and reproducible. However, interobserver agreement on <sup>68</sup>Ga-FAPI PET/CT interpretations remains poorly understood, and consistent interpretation might be challenging given the heterogeneity of tumor diseases and limited

Received Nov. 23, 2022; revision accepted Feb. 21, 2023.

For correspondence or reprints, contact Riccardo Mei (riccardo.mei@studio.unibo.it).

\*Contributed equally to this work.

Published online May 25, 2023.

COPYRIGHT © 2023 by the Society of Nuclear Medicine and Molecular Imaging.

availability of these novel radiotracers. Regardless of the diagnostic value of an imaging method, the overall performance is closely linked to interobserver agreement and variability, which need to be established before implementation into clinical routine (11,12). To address this need, we prospectively evaluated interobserver agreement and accuracy for  $^{68}\text{Ga}$ -FAPI PET/CT interpretations in different tumor entities and compared findings among readers with various levels of experience.

## MATERIALS AND METHODS

### Study Design and Patients

Fifty patients who underwent  $^{68}\text{Ga}$ -FAPI-46 PET/CT for imaging of the following tumor types were retrospectively selected from 2 institutional databases (University Hospital Essen and University Hospital Bologna): sarcoma ( $n = 10$ ), pancreatic adenocarcinoma ( $n = 10$ ), colorectal cancer ( $n = 10$ ), genitourinary cancer ( $n = 10$ ) and miscellaneous cases ( $n = 10$ ) consisting of various other cancer entities (breast cancer [ $n = 1$ ]; lung cancer [ $n = 2$ ]; pleural mesothelioma [ $n = 2$ ]; cholangiocarcinoma [ $n = 2$ ]; hepatocellular carcinoma [ $n = 1$ ]; head and neck cancer [ $n = 1$ ]; and lymphoma [ $n = 1$ ]).

PET/CT-positive lesions were defined by consensus during a joint reading session by 2 expert readers, each with more than 500 prior clinical or research  $^{68}\text{Ga}$ -FAPI-46 PET/CT readings. Expert readers had access to all clinical data. Clinical information, histopathology, tumor markers, and follow-up imaging (CT/MRI or PET/CT) were used as the standard of reference (SOR) for lesion validation. Cases were selected to represent clinical routine, ranging from negative cases ( $n = 8$ , 16%) to extensive disease ( $n = 28$ , 56%), with typical pitfalls. Data were analyzed as part of the prospective interobserver study (NCT04990882) and approved by the local Ethics Committee (permits 19-8991-BO and 20-9485-BO). Patients gave written informed consent to undergo clinical  $^{68}\text{Ga}$ -FAPI-46 PET/CT, and the respective institutional review boards waived individual patient consent for anonymized assessment of their datasets as part of the interobserver study.

### Image Acquisition and Reconstruction

Patient preparation and image acquisition were performed as previously described (13).

In brief, PET was performed on a PET/CT system (Biograph mCT or Vision; Siemens). Scanning was performed at a mean ( $\pm$ SD) of  $14 \pm 9$  min after injection (minimum, 10 min; maximum, 48 min). The injected activity of  $^{68}\text{Ga}$ -FAPI was  $150 \pm 38$  MBq. All PET images were iteratively reconstructed (Vision: 4 iterations, 5 subsets,  $220 \times 220$  matrix, 5-mm Gauss filtering; mCT: 3 iterations, 21 subsets;  $200 \times 200$  matrix, 4-mm Gauss filtering) with time-of-flight information, using the manufacturer's dedicated software (syngo.via for MI; Siemens Healthineers). In all patients, a low-dose CT scan was acquired for attenuation correction (30 mAs, 120 keV,  $512 \times 512$  matrix, 3-mm slice thickness). If available, diagnostic CT scans (with or without contrast enhancement, within 2 wk of  $^{68}\text{Ga}$ -FAPI PET/CT) were provided for the respective case.

### Observers

Fifteen nuclear medicine physicians (5 with additional radiology training) who had at least 1 y of experience with other tracers ( $^{18}\text{F}$ -FDG,  $^{18}\text{F}/^{68}\text{Ga}$ -PSMA,  $^{68}\text{Ga}$ -DOTA-peptides) were prospectively recruited as observers. They were from 10 centers in Europe ( $n = 12$ ), North America ( $n = 2$ ), and Australia ( $n = 1$ ). On the basis of the previously reported number of  $^{68}\text{Ga}$ -FAPI PET/CT interpretations, observers were classified as having a low (<30 prior  $^{68}\text{Ga}$ -FAPI PET/CT studies;  $n = 5$ ), intermediate (30–300 studies;  $n = 5$ , 4 of whom also had radiology education), or high level of experience (>300 studies;  $n = 5$ , 1 of whom also had radiology education). The observers

reviewed all provided  $^{68}\text{Ga}$ -FAPI PET/CT datasets ( $n = 50$ ). Each dataset included low-dose CT and attenuation-corrected PET images and, if available, diagnostic CT images.

### Guidelines for Visual Interpretation

For standardized visual interpretation of datasets, the observers were provided a written guide (Supplemental Data File 1; supplemental materials are available at <http://jnm.snmjournals.org>), 4 teaching cases (Supplemental Data File 2), an electronic case report form, and 1 test patient dataset with disclosed data entries. Furthermore, the observers were asked to learn about  $^{68}\text{Ga}$ -FAPI-46 PET/CT pitfalls (13).

Only necessary patient information was disclosed to the observers before image interpretation: indication for  $^{68}\text{Ga}$ -FAPI PET/CT (primary diagnosis, prior surgery/therapy, staging or restaging of metastatic disease), age (y), weight (kg), injected dose (MBq), and uptake time (min). The observers were masked to all other clinical data. Visual image interpretation for the presence or absence of malignant disease was reported for predefined organ and region categories.

### Statistical Analyses and Reference Standard

Agreement among observer groups was evaluated using the Fleiss  $\kappa$  (14).

Ninety-five percent CIs are reported for  $\kappa$ . The interpretation of  $\kappa$  was based on a classification provided by Landis and Koch (15): 0.0, poor; 0.0–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, almost-perfect reproducibility. SOR was defined as the decision of the experienced expert readers, unmasked to all clinical information (e.g., histopathology, follow-up imaging, and tumor markers).

Overall agreement, defined as the total agreement of an observer on all categories (primary tumor, local nodal involvement, distant metastasis), and sensitivity, specificity, and positive and negative predictive value compared with SOR were calculated. The difference between 2 groups was assessed by the Student  $t$  test at a significance level of  $P$  less than 0.05. Statistical analyses were performed using SPSS software (version 26.0, SPSS Inc.) for all other statistical analyses, and graphs were generated using GraphPad Prism (version 9.1.0; GraphPad Software). At least substantial agreement ( $\kappa \geq 0.6$ ) on visual interpretation of all scans for the 3 major staging categories (primary tumor, local nodal involvement, distant metastasis) and diagnostic accuracy of at least 80% were defined as acceptable performance.

## RESULTS

### Patient Characteristics

Table 1 summarizes the patient characteristics. Twenty-seven male (54.0%) and 23 female (46.0%) patients were included in this study, with a mean age of  $58 \pm 18$  y (range, 19–83 y). Twenty-one patients were imaged at initial diagnosis (42.0%), and 29 patients (58.0%) underwent  $^{68}\text{Ga}$ -FAPI PET/CT for disease restaging. Diagnostic CT was available for 68% of the patients (34 patients: 9 with sarcoma, 9 with pancreatic adenocarcinoma, 5 with colorectal cancer, 4 with genitourinary cancer, and 7 with miscellaneous cancers), and low-dose CT was performed on 16 (32%) patients.  $^{68}\text{Ga}$ -FAPI PET/CT studies were interpreted as positive for local or metastatic cancer lesions in 42 of 50 (84.0%) patients by the reference readers: primary tumor was present in 10 patients (20.0%); 4 patients (8.0%) had lymph node-positive disease, whereas 28 (56.0%) were staged as bone-positive or organ-positive (distant metastasis), respectively.

### Image Interpretation: Interobserver Agreement

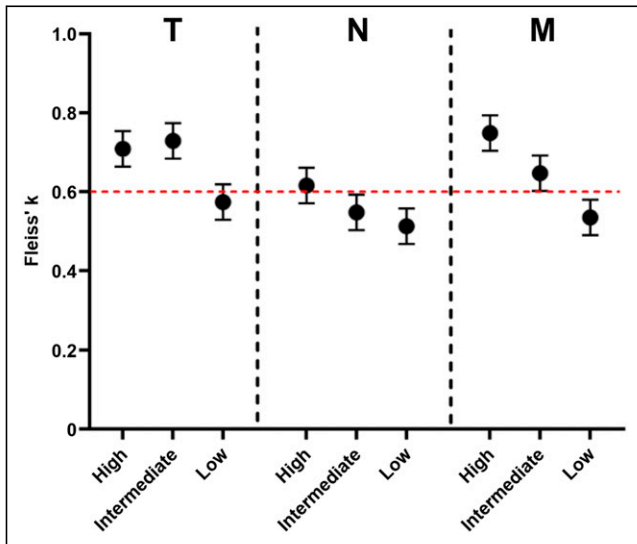
The interobserver agreement on visual image interpretation is summarized in Figure 1 and Table 2. Highly experienced observers agreed substantially on all categories (primary tumor,  $\kappa = 0.71$ ;

**TABLE 1**  
Patient Characteristics (n = 50)

Characteristic	Data
<b>Age (y)</b>	
Mean	58 (SD, 16)
Median	61 (range, 19, 83)
<b>Sex</b>	
Female	23 (46.0%)
Male	27 (54.0%)
<b>Indication</b>	
Initial staging	21 (42.0%)
Restaging	29 (58.0%)
<b>Tumor stage</b>	
No tumor	8 (16.0%)
Local	10 (20.0%)
Local nodal metastatic	4 (8.0%)
Metastatic (organ and bone)	28 (56.0%)

Data are number followed by percentage in parentheses, except for age.

local nodal involvement,  $\kappa = 0.62$ ; distant metastasis,  $\kappa = 0.75$ ), whereas observers with intermediate experience showed substantial agreement on primary tumor ( $\kappa = 0.73$ ) and distant metastasis ( $\kappa = 0.65$ ) and moderate agreement on nodal involvement ( $\kappa = 0.55$ ). Observers with low experience had moderate agreement on all categories (primary tumor,  $\kappa = 0.57$ ; local nodal involvement,  $\kappa = 0.51$ ; distant metastasis,  $\kappa = 0.54$ ). Regardless of the experience level, the overall interobserver agreement was moderate for primary tumor ( $\kappa = 0.56$ ) and local nodal involvement ( $\kappa = 0.55$ ) but substantial for other metastatic lesions ( $\kappa = 0.69$ ).



**FIGURE 1.** Interobserver agreement by primary tumor (T), local nodal involvement (N), distant metastasis (M), and experience level (high, intermediate, and low).

**TABLE 2**

Interobserver Agreement on Visual Image Interpretation

Category	Agreement level	Mean $\kappa$	Inferior 95% CI	Superior 95% CI
T	High	0.709	0.705	0.710
	Intermediate	0.729	0.727	0.732
	Low	0.574	0.572	0.577
N	Any	0.563	0.490	0.633
	High	0.616	0.613	0.619
	Intermediate	0.548	0.546	0.551
M	Low	0.513	0.510	0.516
	Any	0.547	0.472	0.616
	High	0.749	0.746	0.752
M	Intermediate	0.647	0.645	0.650
	Low	0.535	0.532	0.538
	Any	0.686	0.614	0.757

**Image Interpretation: Comparison to SOR**

Table 3 summarizes—by category—sensitivity, specificity, positive and negative predictive value, and accuracy for the entire group and separately for observers with a low, intermediate, or high level of experience. Independent of the experience level, all observers reported primary tumor and distant metastasis stages with high sensitivity (89.0% and 91%). Sensitivity was lower for local nodal involvement stages (79.0%). Because of higher rates of false-positive lesions, specificity was lower for primary tumor stages than for local nodal involvement and distant metastasis stages (65% vs. 81% vs. 77%). Five representative patient examples of disagreement among observers are given in Figure 2.

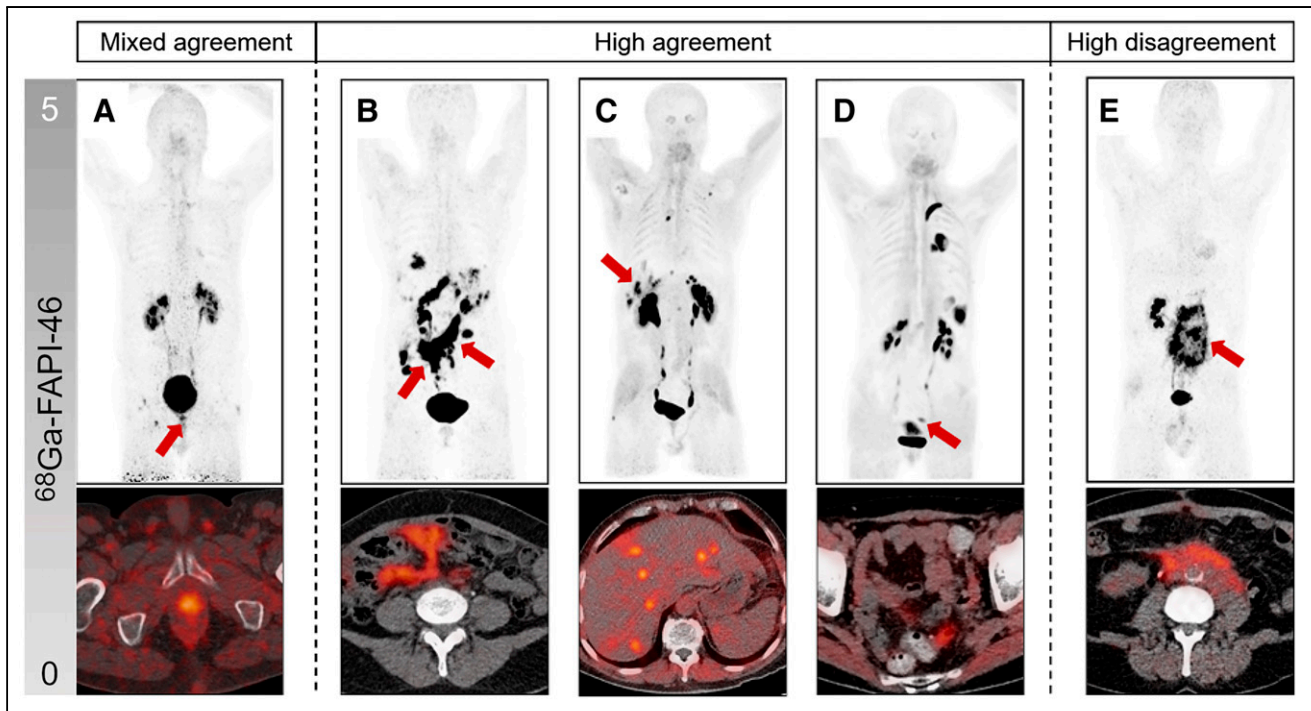
Mean overall agreement with SOR for primary tumor, local nodal involvement, and distant metastasis staging had a  $\kappa$  value

**TABLE 3**

Sensitivity and Specificity for Observer with High, Intermediate, and Low Experience and for All Observers

Category	Agreement level	SE	SP	PPV	NPV	ACC
T	High	0.91	0.63	0.81	0.79	0.81
	Intermediate	0.89	0.73	0.86	0.80	0.84
	Low	0.88	0.59	0.79	0.73	0.77
	All combined	0.89	0.65	0.82	0.77	0.81
N	High	0.77	0.89	0.73	0.91	0.86
	Intermediate	0.83	0.77	0.59	0.92	0.79
	Low	0.88	0.59	0.79	0.73	0.77
	All combined	0.79	0.81	0.61	0.91	0.80
M	High	0.87	0.88	0.91	0.83	0.87
	Intermediate	0.96	0.73	0.83	0.93	0.86
	Low	0.90	0.70	0.80	0.83	0.81
	All combined	0.91	0.77	0.84	0.86	0.85

SE = sensitivity; SP = specificity; PPV = positive predictive value; NPV = negative predictive value; ACC = accuracy.



**FIGURE 2.** Example cases and lesions (arrows) with high agreement, high disagreement, and mixed agreement with SOR. Maximum intensity projections (MIPs) are represented above, and fused transaxial PET/CT images are represented below. (A) Patient with local relapse of urothelial cancer confirmed by histopathology; 7 readers rated positive (2 high, 3 intermediate, 2 low experience), and 8 readers rated lesion negative (3 high, 2 intermediate, 3 low experience). (B) Patient with ovarian cancer; all readers agreed with expert on peritoneal involvement. (C) All readers agreed with expert on liver metastasis in patient with colorectal carcinoma. (D) All readers correctly identified diverticulosis in sarcoma patient. (E) Patient with urothelial cancer and retroperitoneal fibrosis; all readers falsely rated retroperitoneal lesions as malignant and disagreed with expert panel and histopathology.

of 0.64 for the entire group of observers. Observers with high and intermediate experience showed substantial agreement with SOR (high,  $\kappa = 0.69$ ; intermediate,  $\kappa = 0.66$ ), whereas readers with low experience had moderate agreement with SOR ( $\kappa = 0.56$ ) (Table 4). Observers with low and intermediate experience had significantly lower agreement than did highly experienced observers ( $\kappa = 0.56$  and  $0.66$  vs.  $0.69$ ,  $P < 0.001$ ). Irrespective of the stage categories, observers with high and intermediate experience had higher sensitivity, specificity, positive and negative predictive values, and accuracy than did observers with low experience (sensitivity, 87% vs. 91% vs. 74%; specificity, 82% vs. 75% vs. 70%; accuracy, 85% vs. 83% vs. 78%) (Table 4).

The interobserver agreement and accuracy of observers were analyzed separately for the different tumor entities (Table 5). Substantial overall agreement was observed for the miscellaneous

( $\kappa = 0.75$ ), sarcoma ( $\kappa = 0.74$ ), colorectal ( $\kappa = 0.70$ ), and genitourinary cases ( $\kappa = 0.63$ ). Only interpretation of pancreatic cancer cases revealed slight overall agreement ( $\kappa = 0.33$ ). Readers were most sensitive for sarcoma and miscellaneous cases (96% and 97%) and had high accuracy (87% and 89%). Pancreatic cases had the lowest diagnostic performance (sensitivity, 73%; specificity, 61%; accuracy, 66%) (Table 5).

## DISCUSSION

This prospective study involving 50  $^{68}\text{Ga}$ -FAPI46 PET/CT patients with various tumor entities demonstrated substantial reproducibility for highly experienced observers, as well as—with limitations—for observers with intermediate experience.

**TABLE 4**

Overall Sensitivity and Specificity for Observer with High, Intermediate, or Low Experience

Experience level	$\kappa$	SE	SP	PPV	NPV	ACC
High	0.69	0.87	0.82	0.83	0.86	0.85
Intermediate	0.66	0.91	0.75	0.79	0.89	0.83
Low	0.56	0.74	0.70	0.74	0.84	0.78

SE = sensitivity; SP = specificity; PPV = positive predictive value; NPV = negative predictive value; ACC = accuracy.

**TABLE 5**

Agreement and Accuracy per Tumor Entity Irrespective of Reader Experience

Tumor entity	$\kappa$	SE	SP	PPV	NPV	ACC
Colorectal	0.68	0.84	0.85	0.83	0.86	0.85
Genitourinary	0.63	0.63	0.82	0.75	0.87	0.82
Miscellaneous	0.75	0.97	0.73	0.88	0.93	0.89
Pancreas	0.33	0.73	0.61	0.59	0.75	0.66
Sarcoma	0.74	0.96	0.77	0.83	0.95	0.87

SE = sensitivity; SP = specificity; PPV = positive predictive value; NPV = negative predictive value; ACC = accuracy.



On the basis of our predefined criteria, we recommend initial training with at least 300 representative patient cases (high experience) to reach an acceptable diagnostic performance for clinical and research interpretations of  $^{68}\text{Ga}$ -FAPI PET/CT scans. This is comparable to other recent radiotracer training recommendations (16). Training cases should include not just tumor findings (low and extensive tumor burden) but also common pitfalls, such as inflammatory changes or degenerative or posttraumatic bone lesions (13).

In recent years,  $^{68}\text{Ga}$ -FAPI PET/CT has been promoted as a novel highly tumor-specific imaging modality with beneficial radiotracer kinetics and low background uptake (17). Nonetheless, because of the heterogeneity of tumor diseases, variance in fibroblast activation protein expression of tumor and stromal cells, and potential pitfalls, the interpretation of  $^{68}\text{Ga}$ -FAPI PET/CT scans might prove more challenging than expected (8,9,13,18–22). Sources of misinterpretations and false-positive ratings of lesions include a variety of pitfalls similar to those for  $^{18}\text{F}$ -FDG but differing from those for other recent targeted radioligands such as  $^{68}\text{Ga}$ -PSMA (23,24). Examples of these pitfalls include inflammatory uptake (e.g., pancreatitis and myocarditis), degenerative bone lesions, and benign tumors (8,13,25). Because  $^{68}\text{Ga}$ -FAPI is relatively new, the list of pitfalls and benign findings is still evolving and therefore must be constantly updated. Radiotracers such as  $^{68}\text{Ga}$ -PSMA for prostate cancer or  $^{68}\text{Ga}$ -DOTATATE/DOTATOC for neuroendocrine tumors are entity-specific. On the other hand,  $^{68}\text{Ga}$ -FAPI is not specific for any malignancy and can be used for a variety of cancers. This might explain why physicians need to be trained on a high number of cases (300) for  $^{68}\text{Ga}$ -FAPI PET to achieve substantial reproducibility, compared with less than 50 patient cases for  $^{68}\text{Ga}$ -PSMA and  $^{68}\text{Ga}$ -DOTATATE/DOTATOC PET (16,26). Our data underline the necessity of high reader experience for best  $^{68}\text{Ga}$ -FAPI PET agreement and diagnostic performance, although intermediate-level readers do show appropriate results as well. As compared with our study on the  $^{68}\text{Ga}$ -FAPI tracer, other interobserver agreement studies on  $^{68}\text{Ga}$ -PSMA and  $^{68}\text{Ga}$ -DOTATATE PET/CT reported higher reproducibility with lower numbers of recommended initial training cases. This difference is partly due to higher specificity and tumor signal given by receptor radiotracers.

Our cohort consisted of a high proportion of patients with metastatic disease, which reflects the likely clinical-use scenario for  $^{68}\text{Ga}$ -FAPI PET/CT as a whole-body staging tool. We observed that despite high sensitivity for local staging, specificity was only 63% even for highly experienced readers. This is likely due to high nonmalignant tracer uptake after surgery or tracer uptake due to inflammatory changes, both frequently noted in pancreatic cancer patients in our cohort and likely reflecting chronic pancreatitis or scarred tissue after pancreatectomy (8,27). A careful review of the patient's history before image interpretation, and sufficient experience, will therefore be required to avoid a negative impact on clinical management.

This study comes with limitations. Because  $^{68}\text{Ga}$ -FAPI PET/CT is applied for different indications, we decided on a balanced oncologic patient cohort from 2 centers. Only a small proportion of tumor entities was reflected by the patient cohort, and a larger number of patients would be necessary to draw conclusions on agreement and diagnostic performance for distinct tumor entities or imaging of benign disease. Although previous interobserver studies were focused on a single tumor entity (i.e., neuroendocrine tumors for  $^{68}\text{Ga}$ -DOTATATE and prostate cancer for  $^{68}\text{Ga}$ -PSMA) (16,26),  $^{68}\text{Ga}$ -FAPI could be used for a variety of malignancies. For this purpose, we aimed to include representative patients with different tumor entities. We acknowledge that

only one fifth of the cases presented with primary tumor and that PET/CT was performed after the primary treatment in almost 60% of the patients. However, we consider restaging of disease to be a representative scenario to test the reproducibility of  $^{68}\text{Ga}$ -FAPI PET interpretation among readers with various levels of experience. The skill of a reader is determined by multiple factors, including clinical knowledge and general experience in imaging and may vary with the oncologic focus of the reader or center (e.g., lung cancer vs. breast cancer vs. sarcoma). We tried to address this limitation by recruitment of readers worldwide and by inclusion of different tumor entities.

## CONCLUSION

Before clinical implementation of  $^{68}\text{Ga}$ -FAPI PET/CT, we recommend a high experience level based on at least 300 training cases for substantial agreement and diagnostic performance in all categories (primary tumor, local nodal involvement, and distant metastasis). On the basis of the different clinical scenarios and pitfalls, more extensive training is required for  $^{68}\text{Ga}$ -FAPI PET/CT than for other radiotracers.

## DISCLOSURE

Lukas Kessler is a consultant for BTG and AAA and received fees from Sanofi outside the submitted work. Manuel Weber reports personal fees from Boston Scientific, Terumo, Advanced Accelerator Applications, and Eli Lilly. Wolfgang Fendler reports fees from SOFIE Biosciences (research funding), Janssen (consultant, speakers' bureau), Calyx (consultant), Bayer (consultant, speakers' bureau, research funding), Parexel (image review), and AAA (speakers' bureau) outside the submitted work. Benedikt Schaarschmidt received a research grant from PharmaCept for an undergoing investigator-initiated study not related to this paper. No other potential conflict of interest relevant to this article was reported.

## KEY POINTS

**QUESTION:** To what extent do  $^{68}\text{Ga}$ -FAPI PET/CT interpretations agree on various tumor entities for observers with different levels of experience?

**PERTINENT FINDINGS:** We observed substantial overall agreement and high accuracy for observers with a high experience level and also, partially, for intermediate-level observers.

**IMPLICATIONS FOR PATIENT CARE:** Before clinical implementation of  $^{68}\text{Ga}$ -FAPI PET/CT, we recommend a high level of reader experience based on at least 300 training cases.

## REFERENCES

1. Puré E, Blomberg R. Pro-tumorigenic roles of fibroblast activation protein in cancer: back to the basics. *Oncogene*. 2018;37:4343–4357.
2. Loktev A, Lindner T, Burger EM, et al. Development of fibroblast activation protein-targeted radiotracers with improved tumor retention. *J Nucl Med*. 2019;60:1421–1429.
3. Lindner T, Loktev A, Altmann A, et al. Development of quinoline-based theranostic ligands for the targeting of fibroblast activation protein. *J Nucl Med*. 2018;59:1415–1422.
4. Ferdinandus J, Frago Costa P, Kessler L, et al. Initial clinical experience with  $^{90}\text{Y}$ -FAPI-46 radioligand therapy for advanced stage solid tumors: a case series of nine patients. *J Nucl Med*. 2022;63:727–734.
5. Fendler WP, Pabst KM, Kessler L, et al. Safety and efficacy of  $^{90}\text{Y}$ -FAPI-46 radioligand therapy in patients with advanced sarcoma and other cancer entities. *Clin Cancer Res*. 2022;28:4346–4353.

6. Giesel FL, Kratochwil C, Schlittenhardt J, et al. Head-to-head intra-individual comparison of biodistribution and tumor uptake of  $^{68}\text{Ga}$ -FAPI and  $^{18}\text{F}$ -FDG PET/CT in cancer patients. *Eur J Nucl Med Mol Imaging*. 2021;48:4377–4385.
7. Kessler L, Ferdinandus J, Hirmas N, et al.  $^{68}\text{Ga}$ -FAPI as a diagnostic tool in sarcoma: data from the  $^{68}\text{Ga}$ -FAPI PET prospective observational trial. *J Nucl Med*. 2022;63:89–95.
8. Röhrich M, Naumann P, Giesel FL, et al. Impact of  $^{68}\text{Ga}$ -FAPI PET/CT imaging on the therapeutic management of primary and recurrent pancreatic ductal adenocarcinomas. *J Nucl Med*. 2021;62:779–786.
9. Dendl K, Koerber SA, Finck R, et al.  $^{68}\text{Ga}$ -FAPI-PET/CT in patients with various gynecological malignancies. *Eur J Nucl Med Mol Imaging*. 2021;48:4089–4100.
10. Syed M, Flechsig P, Liermann J, et al. Fibroblast activation protein inhibitor (FAPI) PET for diagnostics and advanced targeted radiotherapy in head and neck cancers. *Eur J Nucl Med Mol Imaging*. 2020;47:2836–2845.
11. Bankier AA, Levine D, Halpern EF, Kressel HY. Consensus interpretation in imaging research: is there a better way? *Radiology*. 2010;257:14–17.
12. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research: getting better but still not good. *JAMA*. 1995;274:645–651.
13. Kessler L, Ferdinandus J, Hirmas N, et al. Pitfalls and common findings in  $^{68}\text{Ga}$ -FAPI-PET: a pictorial analysis. *J Nucl Med*. 2022;63:890–896.
14. Hale CA, Fleiss JL. Interval estimation under two study designs for kappa with binary classifications. *Biometrics*. 1993;49:523–534.
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
16. Fendler WP, Calais J, Allen-Auerbach M, et al.  $^{68}\text{Ga}$ -PSMA-11 PET/CT interobserver agreement for prostate cancer assessments: an international multicenter prospective study. *J Nucl Med*. 2017;58:1617–1623.
17. Giesel FL, Kratochwil C, Lindner T, et al.  $^{68}\text{Ga}$ -FAPI PET/CT: biodistribution and preliminary dosimetry estimate of 2 DOTA-containing FAP-targeting agents in patients with various cancers. *J Nucl Med*. 2019;60:386–392.
18. Dendl K, Koerber SA, Adeberg S, et al. Physiological FAP-activation in a postpartum woman observed in oncological FAPI-PET/CT. *Eur J Nucl Med Mol Imaging*. 2021;48:2059–2061.
19. Gündoğan C, Güzel Y, Can C, Alabalik U, Kömek H. False-positive  $^{68}\text{Ga}$ -fibroblast activation protein-specific inhibitor uptake of benign lymphoid tissue in a patient with breast cancer. *Clin Nucl Med*. 2021;46:e433–e435.
20. Liu H, Wang Y, Zhang W, Cai L, Chen Y. Elevated  $^{68}\text{Ga}$ -FAPI activity in splenic hemangioma and pneumonia. *Clin Nucl Med*. 2021;46:694–696.
21. Lin R, Lin Z, Zhang J, Yao S, Miao W. Increased  $^{68}\text{Ga}$ -FAPI-04 uptake in Schmorl node in a patient with gastric cancer. *Clin Nucl Med*. 2021;46:700–702.
22. Kratochwil C, Flechsig P, Lindner T, et al.  $^{68}\text{Ga}$ -FAPI PET/CT: tracer uptake in 28 different kinds of cancer. *J Nucl Med*. 2019;60:801–805.
23. Rischpler C, Beck TI, Okamoto S, et al.  $^{68}\text{Ga}$ -PSMA-HBED-CC uptake in cervical, celiac, and sacral ganglia as an important pitfall in prostate cancer PET imaging. *J Nucl Med*. 2018;59:1406–1411.
24. Ibrahim N, Cho S, Hall L, Perlman S. Is it cancer? Potential pitfalls of FDG PET/CT imaging [abstract]. *J Nucl Med*. 2016;57(suppl 2):1315.
25. Zhang X, Song W, Qin C, Liu F, Lan X. Non-malignant findings of focal  $^{68}\text{Ga}$ -FAPI-04 uptake in pancreas. *Eur J Nucl Med Mol Imaging*. 2021;48:2635–2641.
26. Fendler WP, Barrio M, Spick C, et al.  $^{68}\text{Ga}$ -DOTATATE PET/CT interobserver agreement for neuroendocrine tumor assessment: results of a prospective study on 50 patients. *J Nucl Med*. 2017;58:307–311.
27. Ferdinandus J, Kessler L, Hirmas N, et al. Equivalent tumor detection for early and late FAPI-46 PET acquisition. *Eur J Nucl Med Mol Imaging*. 2021;48:3221–3227.