Does whole-body Patlak 18F-FDG PET imaging improve lesion detectability in clinical oncology?

FAHRNI, Guillaume, et al.

Abstract

Single-pass whole-body (WB) 18F-FDG PET/CT imaging is routinely employed for the clinical assessment of malignant, infectious, and inflammatory diseases. Our aim in this study is the systematic clinical assessment of lesion detectability in multi-pass WB parametric imaging enabling direct imaging of the highly quantitative 18F-FDG influx rate constant Ki, as a complement to standard-of-care standardized uptake value (SUV) imaging for a range of oncologic studies.

Reference

FAHRNI, Guillaume, et al. Does whole-body Patlak 18F-FDG PET imaging improve lesion detectability in clinical oncology? European Radiology, 2019

DOI : 10.1007/s00330-018-5966-1
PMID : 30689031

Available at:
http://archive-ouverte.unige.ch/unige:113671

Disclaimer: layout of this document may differ from the published version.
Does whole-body Patlak $^{18}$F-FDG PET imaging improve lesion detectability in clinical oncology?

Guillaume Fahrni$^1$ · Nicolas A. Karakatsanis$^2$ · Giulia Di Domenicantonio$^1$ · Valentina Garibotto$^{1,3}$ · Habib Zaidi$^{1,4,5,6}$

Received: 27 August 2018 / Revised: 7 November 2018 / Accepted: 11 December 2018  © European Society of Radiology 2019

Abstract

Purpose

Single-pass whole-body (WB) $^{18}$F-FDG PET/CT imaging is routinely employed for the clinical assessment of malignant, infectious, and inflammatory diseases. Our aim in this study is the systematic clinical assessment of lesion detectability in multi-pass WB parametric imaging enabling direct imaging of the highly quantitative $^{18}$F-FDG influx rate constant $K_i$, as a complement to standard-of-care standardized uptake value (SUV) imaging for a range of oncologic studies.

Methods

We compared SUV and $K_i$ images of 18 clinical studies of different oncologic indications (lesion characterization and staging) including standard-of-care SUV and dynamic WB PET protocols in a single session. The comparison involved both the visual assessment and the quantitative evaluation of SUV mean, SUV max, $K_i$ mean, $K_i$ max, tumor-to-background ratio (TBR SUV, TBR $K_i$), and contrast-to-noise ratio (CNR SUV, CNR $K_i$) quality metrics.

Results

Overall, both methods provided good-quality images suitable for visual interpretation. A total of 118 lesions were detected, including 40 malignant (proven) and 78 malignant (unproven) lesions. Of those, 111 were detected on SUV and 108 on $K_i$ images. One proven malignant lesion was detected only on $K_i$ images whereas none of the proven malignant lesions was visible only on SUV images. The proven malignant lesions had overall higher $K_i$ TBR and CNR scores. One unproven lesion, which was later confirmed as benign, was detected only on the SUV images (false-positive). Overall, our results from 40 proven malignant lesions suggested improved sensitivity (from 92.5 to 95%) and accuracy (from 90.24 to 95.12%) and potentially enhanced specificity with $K_i$ over SUV imaging.

Conclusion

Oncologic WB Patlak $K_i$ imaging may achieve equivalent or superior lesion detectability with reduced false-positive rates when complementing standard-of-care SUV imaging.

Key Points

- The whole-body spatio-temporal distribution of $^{18}$F-FDG uptake may reveal clinically useful information on oncologic diseases to complement the standard-of-care SUV metric.
- Parametric imaging resulted in less false-positive indications of non-specific $^{18}$F-FDG uptake relative to SUV.
- Parametric imaging may achieve equivalent or superior $^{18}$F-FDG lesion detectability than standard-of-care SUV imaging in oncology.

Keywords  Positron emission tomography · Molecular imaging · Tumors

Electronic supplementary material

The online version of this article (https://doi.org/10.1007/s00330-018-5966-1) contains supplementary material, which is available to authorized users.

Nicolas A. Karakatsanis
nak2032@med.cornell.edu

Habib Zaidi
habib.zaidi@hcuge.ch

1 Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland

2 Division of Radiopharmaceutical Sciences, Department of Radiology, Weill Cornell Medical College of Cornell University, New York, NY 10021, USA

3 Faculty of Medicine, University of Geneva, 1205 Geneva, Switzerland

4 Geneva University Neurocenter, University of Geneva, 1205 Geneva, Switzerland

5 Department of Nuclear Medicine and Molecular Imaging, University of Groningen, 9700 RB Groningen, Netherlands

6 Department of Nuclear Medicine, University of Southern Denmark, DK-500 Odense, Denmark

Published online: 28 January 2019
Abbreviations

- $^{18}$F-FDG: $^{18}$F-Fluorodeoxyglucose
- CT: Computed tomography
- FOV: Field-of-view
- IDIF: Image-derived input function
- LV: Left ventricle
- PET: Positron emission tomography
- PET/CT: Positron emission tomography/computed tomography
- ROI: Region of interest
- SUV: Standardized uptake value
- TOF: Time-of-flight
- VOI: Volume of interest
- WB: Whole-body

Introduction

Whole-body (WB) $^{18}$F-FDG PET/CT imaging is widely used in clinical oncology for multiple indications, including lesion characterization, staging, restaging, and therapy monitoring [1, 2]. In current clinical practice, oncologic PET/CT exams mainly consist of scans performed across multiple bed positions over a single time period or frame at late time points, when equilibrium is reached. The acquired data are then averaged over the acquisition time frame and reconstructed to produce conventional (static) PET images. The images are quantified in units of standardized uptake value (SUV), which serve as a surrogate of static metabolic activity normalized to the body weight and the administered tracer dosage [3]. Yet, SUV remains a semi-quantitative index subject to a large variability due to several factors related to the scanning procedure or the patient’s physiology, such as patient’s body composition, plasma glucose level, time between injection and measurement, and duration of the scan time frame [4]. Moreover, static imaging is unable to provide information about the kinetics of $^{18}$F-FDG in the regions of interest [5].

The dynamic course of the $^{18}$F-FDG spatial distribution in the targeted tissues may reveal highly useful clinical information on tissue’s metabolic properties, such as the metabolic rate of glucose uptake post $^{18}$F-FDG injection. These image metrics could, in turn, facilitate tumor characterization and therapy response assessment [6–9]. For this reason, dynamic PET imaging techniques were also introduced allowing the scanning of a limited axial field-of-view (single bed position) over time to enable the extraction of important tracer kinetic parameters, such as the uptake rate. The tracer kinetic parameters were initially estimated on a region level by averaging the dynamic PET images within specific volumes of interest (VOIs) [10, 11]. Later, larger computational resources allowed for voxel-wise estimation of the tracer kinetic parameters at the cost of higher noise levels, thus leading to parametric imaging [12]. Despite the potential for better lesion detectability and reduced rate of false positives [13], dynamic and parametric imaging has hitherto been limited to research setting and single-bed field-of-views (FOVs) [14–16].

Different strategies were developed for the generation of parametric images, ranging from streamlined graphical analysis, such as the Patlak method [17], to spectral analysis [18] and to the most complex full compartmental modeling methods [19]. Although the two latter techniques are more informative, they can be unstable, sensitive to statistical noise, and difficult to adopt in the clinic, due to the large number of fitting parameters and the high order of the associated kinetic models. In this regard, graphical analysis can be a simpler, yet robust, alternative approach enabling the direct estimation of the primary kinetic macro-components of the tracer uptake across multiple-bed FOVs.

However, WB dynamic PET acquisitions present the additional challenge of segmenting a given dynamic scan period across multiple bed positions, thereby inevitably introducing time gaps in between the time frames, which reduce the signal-to-noise ratio [20]. Nevertheless, recent technological advances in commercial clinical PET scanners, such as the use of longer axial FOVs [21], time-of-flight (TOF) information [22], and resolution modeling [23], allowed significant enhancement of the noise-equivalent count rate and sensitivity. In fact, the first introduced framework of WB dynamic $^{18}$F-FDG PET/CT clinical imaging involved a 6-min dynamic acquisition over the heart bed, followed by four or six dynamic PET passes across multiple beds, and was successfully applied on a pilot cohort of oncology cases demonstrating its clinical feasibility [24–27]. The utilization of just two time-frame acquisitions (passes) per bed position has also been reported [28]. Nevertheless, the use of six WB passes demonstrated improved performance, in terms of noise and bias, for the WB parametric images to enable the combined estimation of SUV and multi-parametric PET images in a single scan session [29].

In this work, we systematically compare lesion detectability between Patlak and SUV $^{18}$F-FDG WB clinical PET images, as extracted from the same cohort of oncologic patients. Our ultimate aim is to assess any benefits in WB $^{18}$F-FDG lesion detectability by utilizing Patlak-based glucose metabolism metrics.

Materials and methods

Patient population

This is a prospective pilot study to facilitate the development of a noninvasive quantitative assessment strategy for dynamic whole-body $^{18}$F-FDG oncologic PET imaging. We included adult patients who were referred for staging or restaging of lung or abdominal lesions by $^{18}$F-FDG PET/CT. The
institutional ethics committee approved this study and all patients gave informed consent.

**Data acquisition and image reconstruction**

PET/CT scans were performed on a Siemens Biograph™ mCT scanner following injection of a standard dose of 3.5 MBq/kg (3.71 ± 1.05 MBq/kg) of $^{18}$F-FDG. The patients underwent a scanning protocol of about 80 min consisting of the following (Fig. 1): (i) a low-dose CT for attenuation correction; (ii) a 6-min dynamic single-bed PET acquisition centered in the heart region for extrapolation of the input function (IF), starting simultaneously with the $^{18}$F-FDG injection; (iii) a subsequent set of 13 whole-body PET scans in continuous bed motion (CBM) of about 3 min each; (iv) a subsequent SUV WB CBM scan of about 20 min used as a reference for comparison was acquired at the standard-of-care scan time window for $^{18}$F-FDG studies (starting 60 min post-injection); and (v) a contrast-enhanced CT scan (13/18 patients) or a non-enhanced CT scan when contraindicated (i.e., renal failure).

All images were reconstructed using 3D iterative OSEM algorithm and parameters used in the clinic (2 iterations and 21 subsets), including TOF and resolution modeling [30], with a post-reconstruction Gaussian filter of 2 mm FWHM.

**Methodology for whole-body parametric imaging**

The Patlak graphical analysis utilized by the employed multipass WB parametric imaging framework is based on an irreversible two-tissue compartment kinetic model: $^{18}$F-FDG in the blood plasma is reversibly absorbed (intra-cellular) by specific tissues in a free non-metabolized state (first tissue compartment) and then, unlike standard glucose, is irreversibly metabolized or phosphorylated (second compartment) into $^{18}$F-FDG6P [31]. The exchanges between the compartments are modeled by the micro-parameters $K_1$ (plasma to non-phosphorylated compartment), $k_2$ (non-phosphorylated to plasma compartment), and $k_3$ (non-phosphorylated to phosphorylated compartment) rate constants. A zero efflux (dephosphorylation) rate constant ($k_4 = 0$) between the phosphorylated and the non-phosphorylated compartments is assumed. FDG $k_4$ parameter is widely assumed to be negligible and can be considered to be zero for the majority of tissues ($k_4$ may be positive in some regions, such as the liver, but its effect on the data is negligible for the early uptake phase). All the exchanges between the different compartments can be resumed in the $^{18}$F-FDG macro-parameter of $K_i$, denoting the net uptake or influx rate constant:

$$K_i = \frac{K_1 k_3}{k_2 + k_3}$$  \hspace{1cm} (1)
In the assumption of irreversible trapping of the tracer, $K_i$ can be directly estimated using the standard Patlak analysis on a voxel-by-voxel basis, resulting in $K_i$ parametric maps [17]. We generated such $K_i$ images by implementing an in-house MATLAB code (MathWorks Inc.). The tissue activity, $C(t)$, is linked to blood plasma time-activity concentration or input function (IF), $C_p(t)$, via the standard Patlak model linear equation:

$$C(t) = K_i C_p(t) \, dt + V C_p(t)$$  \hspace{1cm} (2)$$

where $K_i$ and $V$ are the uptake rate defined in Eq. 1 and the total blood plasma distribution volume, respectively. The activity in each tissue voxel $C(t)$ and the $C_p(t)$, as measured over time $t$ from the dynamic PET images, are then fitted to the Patlak model equation to estimate via ordinary least squares regression the $K_i$ and $V$ parameters at every voxel.

In this protocol, $C(t)$ is sampled 13 times (once per pass) for each voxel; the $C_p(t)$ is determined using both the initial dynamic acquisition and the 13 passes. Finally, the parametric $K_i$ images are normalized to the maximum SUV of the respective static scan to display the images on the same scale, thus facilitating head-to-head clinical interpretation and image comparison.

**Derivation of the input function**

The image-derived IF (IDIF) $C_p(t)$ is obtained from a region of interest (ROI) drawn in the left ventricle (LV) of the heart on images from both scans (ii) and (iii): the former was sampled in 12 frames of 10 s followed by 12 frames of 20 s, while a single IDIF value was obtained from the cardiac bed frame of each of the 13 WB passes.

A number of studies reported that the IDIF can also be extracted from the aorta rather than the left ventricle [32]. Therefore, we evaluated the difference between LV and aortic IDIF for two patients where arterial blood sampling was also performed as the gold standard IF. Comparable LV and aortic IDIFs were found, although both underestimated the gold standard IF at the peak (Fig. 2, left panel) and overestimated it at later times. However, the integral of the IF (Fig. 2, right panel), which is the only IF metric considered in Patlak analysis, varied by less than 5% in between LV, aortic, and gold standard IF resulting in minimal impact on the derived parametric images.

**Image analysis**

**Visual interpretation** SUV and $K_i$ images were evaluated qualitatively on a dedicated workstation by two physicians (having only access to the clinical indication) who independently analyzed both images in two different and distant sessions followed by consensus reading and the writing of a single report for each set of images summarizing the findings. The overall quality of PET images was rated using a 5-point scale from 1 (inacceptable) to 5 (excellent) on the basis of the subjective impression of the overall quality of the PET data fused on the corresponding CT images, taking into account the impression of smoothness, resolution, sharpness of contours, noise, homogeneity, and presence of artifacts. A single consensus score was assigned by expert physicians. Interobserver evaluation and comparison of different groups of readers are beyond the scope of this work. Both reports were then compared. Image quality was assessed as adequate/inadequate for visual reading and all primary tumors and other distant lesions, such as metastases, lymph nodes, or inflammatory uptake, were identified and quantitatively evaluated. We also included malignant (biopsy proven)/malignant (unproven) lesions that were not visible on either SUV or $K_i$ images, but only on the CT component of the PET/CT examination accounting for the probability that lesions seen on CT but not on PET may be non-malignant. When possible, malignancy was confirmed by histology or other clinical indications provided in patient’s follow-up.

We analyzed separately biopsy-proven or previously known lesions, which were classified as proven malignant, from lesions considered as unproven malignant with high clinical probability for malignancy (i.e., new lesions in a patient with metastases or lesions confirmed on contrast-enhanced CT in case of FDG-non avid lesions, such as hepatocellular carcinoma (HCC)). In cases where a significant number of lesions was identified in the same organ or anatomical region (i.e., multiple abdominal lymph nodes), a maximum of 5 lesions were analyzed to reduce bias.

**Quantitative analysis** A 3D ROI was manually drawn on each lesion to estimate at the same position the SUV$_{\text{max}}$ and $K_i$$_{\text{max}}$ metrics. When the lesions were non-metabolic, the ROIs were drawn on the contrast-enhanced CT images. Background ROI mean, BKG$_{\text{mean}}$, and standard deviation, BKG$_{\text{sd}}$, metrics were extracted from the same positions, between SUV and $K_i$ images, in healthy surrounding tissues. Tumor-to-background ratio (TBR) and contrast-to-noise ratio (CNR) scores were then calculated as defined in Eqs. 3 and 4.

$$\text{TBR} = \frac{(\text{Tumor ROI}_{\text{max}} / \text{Background ROI}_{\text{mean}})-1}{\text{CNR}}$$  \hspace{1cm} (3)$$

Finally, all identified lesions were regrouped in organs regardless of the origin of the lesion (primary tumor, metastases, malignant (unproven) lesion).

**Statistical analysis** Spearman’s rank correlation coefficient ($\rho$) was used to assess the correlation between SUV$_{\text{max}}$ and $K_i$$_{\text{max}}$ of all lesions as...
well as between background values. A p value < 0.01 was considered significant for this test. A Wilcoxon signed-rank test was performed to assess the statistical significance of the differences between TBR$_{K_i}$ and TBR$_{SUV}$ and between CNR$_{K_i}$ and CNR$_{SUV}$ for the 40 malignant lesions. A p value < 0.05 was considered significant for this test.

**Results**

Eighteen patients originally scheduled for an oncologic whole-body $^{18}$F-FDG PET/CT (6 females and 12 males; mean age = 58.8 ± 16.8 years) were enrolled. A description of the cohort, including the indication for the study and the number and location of detected lesions, is reported in Table 1. Both modalities produced good-quality images adequate for visual interpretation. The qualitative inspection of the PET images by expert physicians suggested similar (or lower) image quality for the $K_i$ (4.05), relative to SUV (4.2) images. However, the $K_i$ image quality absolute scores were acceptable, the only discordant case being one patient with significant bulk motion during the long dynamic WB scan. The most notable visual difference, as illustrated in Fig. 3, was the suppression of the blood compartment in $K_i$ images, particularly for organs with a non-negligible blood pool component, such as the liver, the spleen, and the blood vessels (Table 2), resulting in a higher contrast for individual lesions. This is in agreement with previous studies [24–27]. A total of 118 biopsy-proven malignant ($n = 40$) or unproven malignant ($n = 78$) lesions were identified and analyzed.

Of the 40 confirmed malignant lesions, 37 (93%) were detected on SUV and 38 (95%) on $K_i$ images. The malignant lesions have higher $K_i$ TBR ($p < 0.001$) and CNR ($p < 0.001$) scores. The 2 lesions not detected on both modalities were HCC nodules on the same patient (patient no. 18), appearing as non-hypermetabolic. Those two lesions were not detected on $K_i$ images despite their high TBR and CNR scores, which may be attributed to the very low $K_i$ values, as discussed later. The only lesion detected with $K_i$ but not with SUV images was a hepatic lesion (patient no. 1, Fig. 3), a 71-year-old male presenting a suspicious liver mass, described as hypervascular on the contrast-enhanced CT examination. The standard report of the SUV images described the lesion as non-metabolic (SUV$_{max}$ 4.45, BKG$_{mean}$ 2.2, TBR 1.02, CNR 2.82), whereas it was described as hypermetabolic in the report generated from the $K_i$ images ($K_{i,max}$ 1.43, BKG$_{mean}$ 0.24, TBR 5.37, CNR 29.17). Biopsy later confirmed the lesion to be a HCC.

Of the 78 unproven malignant lesions, 74 (95%) were detected with SUV images whereas 70 (90%) of them were detected on $K_i$ images. The four lesions not detected on either modality were non-hypermetabolic liver lesions (patient no. 2 and no. 12) suspected as HCC nodules on the basis of CT and MR imaging, but never proven with a biopsy. The four lesions not detected on $K_i$ images only were the following: a hypermetabolic focal sigmoid activity (patient no. 10); a slightly hypermetabolic lymph node (patient no. 14); a slightly hypermetabolic focal brain activity (patient no. 14); and a lung lesion (patient no. 8, Fig. 4). The first three lesions were not confirmed as either malignant or benign during the follow-up. The last lesion was a false-positive case of a 54-year-old male with a suspicious left lung lesion, confirmed as benign at biopsy. The standard report of the SUV images described the lesion as hypermetabolic (SUV$_{max}$ 1.5, BKG$_{mean}$ 0.29, TBR 4.17, CNR 81.81), whereas it was not detected in the report generated from the $K_i$ images ($K_{i,max}$ 0.47, BKG$_{mean}$ 0.02, TBR 22.5, CNR 2812.5).

Considering the biopsy results as our gold standard outcome for assessing malignancy of lesions, we assessed the sensitivity and accuracy for the 40 lesions proven as malignant. In this group of lesions, the sensitivity in detecting malignancy increased from 37 × 100 40 = 92.5%, in the case of SUV, to 38 × 100 40 = 95%, in the case of $K_i$ imaging. As we only had one biopsy-confirmed negative lesion, we cannot reliably compare the specificity performance between SUV and $K_i$ imaging. Nevertheless, we were able to
correctly assess this single true-negative indication only with $K_i$ imaging, therefore suggesting improved specificity with $K_i$ imaging. As a result, our accuracy in detecting malignancy was further improved from $37 \times 100 = 90.24\%$ with SUV to $39 \times 100 = 95.12\%$ with $K_i$ imaging.

**PET metrics and statistical analysis**

Spearman’s rank correlation coefficient ($\rho$), which was calculated independently of the anatomical region, was 0.777 ($p < 0.001$) for $K_{\text{max}}$-$SUV_{\text{max}}$ of all lesions and 0.414 ($p < 0.001$) for $K_{\text{mean}}$.

Table 1  Clinical characteristics of the patients’ population and summary of the lesions detected

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age</th>
<th>Sex</th>
<th>Malignancy</th>
<th>Detected lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>71</td>
<td>M</td>
<td>Hepatocellular carcinoma</td>
<td>Liver (1)</td>
</tr>
<tr>
<td>2</td>
<td>70</td>
<td>M</td>
<td>Pancreatic adenocarcinoma</td>
<td>Pancreas (1), liver (2), lymph nodes (1)</td>
</tr>
<tr>
<td>3</td>
<td>69</td>
<td>M</td>
<td>Pancreatic adenocarcinoma</td>
<td>Pancreas (1), liver (2), lymph nodes (4)</td>
</tr>
<tr>
<td>4</td>
<td>42</td>
<td>M</td>
<td>Neuroendocrine carcinoma</td>
<td>Lungs (1), subcutaneous tissue (1), lymph nodes (2)</td>
</tr>
<tr>
<td>5</td>
<td>59</td>
<td>M</td>
<td>Hepatocellular carcinoma</td>
<td>Liver (1), lymph nodes (3)</td>
</tr>
<tr>
<td>6</td>
<td>67</td>
<td>M</td>
<td>Lung adenocarcinoma</td>
<td>Lungs (1)</td>
</tr>
<tr>
<td>7</td>
<td>58</td>
<td>M</td>
<td>Rectum adenocarcinoma</td>
<td>Colon (3), lungs (2), subcutaneous tissue (5), peritoneum (5), lymph nodes (3), liver (3), muscles (1)</td>
</tr>
<tr>
<td>8</td>
<td>54</td>
<td>M</td>
<td>Suspicious lung lesion</td>
<td>Lungs (1)</td>
</tr>
<tr>
<td>9</td>
<td>64</td>
<td>M</td>
<td>Melanoma</td>
<td>Lungs (1), liver (2)</td>
</tr>
<tr>
<td>10</td>
<td>56</td>
<td>F</td>
<td>Lung carcinoid</td>
<td>Lungs (1), colon (1)</td>
</tr>
<tr>
<td>11</td>
<td>42</td>
<td>F</td>
<td>Lung adenocarcinoma</td>
<td>Lungs (2), lymph nodes (3)</td>
</tr>
<tr>
<td>12</td>
<td>74</td>
<td>M</td>
<td>Intrahepatic cholangiocarcinoma</td>
<td>Liver (3), lymph nodes (5), brain (1)</td>
</tr>
<tr>
<td>13</td>
<td>45</td>
<td>F</td>
<td>Hepatic adenocarcinoma</td>
<td>Liver (6), lungs (2), colon (2), bones (5), lymph nodes (8)</td>
</tr>
<tr>
<td>14</td>
<td>45</td>
<td>F</td>
<td>Gastric adenocarcinoma</td>
<td>Stomach (1), lymph nodes (3), bones (2), brain (1)</td>
</tr>
<tr>
<td>15</td>
<td>62</td>
<td>F</td>
<td>Lung adenocarcinoma</td>
<td>Lungs (1), colon (1)</td>
</tr>
<tr>
<td>16</td>
<td>59</td>
<td>F</td>
<td>Lung neuroendocrine carcinoma</td>
<td>Lungs (1), adrenal gland (3), bones (1), lymph nodes (4)</td>
</tr>
<tr>
<td>17</td>
<td>60</td>
<td>M</td>
<td>Lung Hodgkin lymphoma</td>
<td>Lungs (1), bones (1), lymph nodes (11)</td>
</tr>
<tr>
<td>18</td>
<td>61</td>
<td>M</td>
<td>Hepatocellular carcinoma</td>
<td>Liver (2), parotid gland (1)</td>
</tr>
</tbody>
</table>
SUV\text{\textsubscript{mean}} of all background regions. Table 3 and Fig. 6 show the correlations for each anatomical region.

TBR\text{\textsubscript{SUV}} values ranged from 0.32 to 59.26, whereas TBR\text{\textsubscript{Ki}} values ranged from 2.6 to 9229. TBR\text{\textsubscript{Ki}} values were higher than TBR\text{\textsubscript{SUV}} values for 38 lesions out of 40 (95%). The Wilcoxon signed-rank test showed statistically significant differences between the two metrics ($p < 0.001$). CNR\text{\textsubscript{SUV}} values ranged from 1.08 to 2693.78, whereas CNR\text{\textsubscript{Ki}} values ranged from 1.82 to 2,307,250. CNR\text{\textsubscript{Ki}} values were higher than CNR\text{\textsubscript{SUV}} values for 33 lesions out of 40 (82.5%). The Wilcoxon signed-rank test showed statistically significant differences between the two metrics ($p < 0.001$).

Multiple findings of muscle focal activity, non-suspicious lymph nodes, digestive metabolism, or other similar foci presumably benign were visible on SUV images but negative on Ki images. As an example of such findings, we report a case of a 45-year-old female (patient no. 14, Supplementary Fig. 1) presenting with a gastric adenocarcinoma. The SUV image report described a non-suspicious thoracic lymph node (SUV\text{\textsubscript{max}} 3.34, BKG\text{\textsubscript{mean}} 2.6, TBR 0.28, CNR 0.31), which was not documented on the Ki image report ($K_{\text{max}}$ 0.73, BKG\text{\textsubscript{mean}} 0.55, TBR 0.33, CNR 0.64).

### Discussion

This study demonstrated that multi-pass whole-body PET $K$-parametric imaging, utilizing robust Patlak graphical analysis, may achieve equivalent or, potentially, superior lesion detectability than standard-of-care SUV imaging with reduced false-positive rates in routine oncology applications. The implemented protocol generated both static SUV and parametric Ki images in a single session, thereby allowing for reliable and objective comparisons between the two modalities. In our study, all produced images were of good quality with all biopsy proven malignant lesions visible in SUV images also depicted in Ki images. However, Ki images proved superior in three situations: one proven malignant lesion only being detected on Ki images, as well as one false-positive lesion and a benign lymph node being positive only in SUV images. A major finding of our work is the

<table>
<thead>
<tr>
<th></th>
<th>SUV range (mean ± SD)</th>
<th>$K$ (normalized) range (mean ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thoracic vertebra</td>
<td>1.10–2.56 (1.88 ± 0.37)</td>
<td>0.11–2.63 (0.70 ± 0.55)</td>
</tr>
<tr>
<td>Lung</td>
<td>0.19–0.76 (0.38 ± 0.16)</td>
<td>4·10$^{-3}$–0.43 (0.12 ± 0.15)</td>
</tr>
<tr>
<td>Blood pool</td>
<td>1.21–2.48 (1.75 ± 0.34)</td>
<td>0–0.45 (0.14 ± 0.16)</td>
</tr>
<tr>
<td>Liver</td>
<td>1.47–3.05 (2.15 ± 0.38)</td>
<td>0–2.88 (0.55 ± 0.70)</td>
</tr>
<tr>
<td>Muscle</td>
<td>0.10–0.91 (0.61 ± 0.17)</td>
<td>0.017–1.40 (0.31 ± 0.33)</td>
</tr>
</tbody>
</table>

---

**Fig. 4** Case of an inflammatory lung lesion (patient no. 8), initially declared as malignant (unproven) in the standard SUV report, confirmed benign by biopsy during the follow-up. This lesion was declared as benign in the $K$ report. a SUV fused with non-contrast abdomen CT, b $K$, fused with non-contrast abdomen CT, c SUV images, and d $K$ images. Note the very different aspect of the arch of the aorta (dashed arrow) on both modalities suggesting suppression of blood pool signal in the $K$ images.

---

**Table 2** Typical SUV and normalized $K$ values for normal tissue in different organs

---

\(\text{Eur Radiol}\)
suppression of the non-specific $^{18}$F-FDG signal in the blood compartments allowing for a higher contrast in abnormal uptake regions, thereby confirming observations made in previous studies [24–27, 33]. This observation suggests increased likelihood of detecting metabolically active lesions surrounded by avid background non-specific uptake, such as in the case of intermediate differentiated HCCs with $^{18}$F-FDG imaging.

The Patlak graphical approach is of low linear complexity, thus allowing for fast computation and wide adoptability of WB parametric PET imaging in clinical practice. On the other hand, it assumes an irreversible $^{18}$F-FDG uptake, a condition not always satisfied in normal liver, kidneys, and certain tumors, such as HCCs [34], thereby leading to $K_i$ underestimation. Thus, an extended non-linear generalized Patlak model, considering a mild degree of reversibility, had been previously proposed [26] and could be easily integrated in our analysis. However, the use of a standard linear graphical analysis model is more preferable, thanks to its robustness to data noise, when the kinetic parameters are indirectly estimated from the dynamic PET images, and higher lesion contrast, when $K_i$ is underestimated only in the background.

We found a strong positive correlation ($\rho = 0.791$) between $K_{i\text{max}}$ and SUV$_{\text{max}}$, and even stronger correlations when looking separately at different organs. This suggests that $K_i$ images will identify hypermetabolic lesions if these are depicted on SUV images. In background neighboring regions, however, we found weak positive correlations ($\rho = 0.392$) between BKG$_{\text{mean}}$-$K_i$ and BKG$_{\text{mean}}$-SUV, which is expected owing to blood pool suppression within $K_i$ images.

The TBR and CNR are almost always higher in $K_i$ images (Fig. 5). On a purely visual examination, most of the malignant lesions were depicted on both modalities, except two HCC nodules appearing as non-hypermetabolic on both modalities and one HCC only depicted on $K_i$ images. Thus, in the majority of cases, the TBR$_{\text{SUV}}$ was largely sufficient to detect the lesions.

In the case of the lung lesions, for example, where the highest TBR values were found, we observe that most lung lesions are associated with a very low metabolism, resulting in lung lesions being easily detected. Overall, a TBR value between 3 and 5 is generally adequate in differentiating the lesion from its background, if the lesion has sufficient metabolic activity. Therefore, improving the TBR beyond this range with parametric imaging will not further improve lesion detectability, although it may improve quantification. On the other hand, detectability may be challenging for lesions with low TBR$_{\text{SUV}}$ scores especially within a highly vascularized organ of high BKG$_{\text{mean}}$ uptake, such as the liver. Detectability may then be improved by the higher TBR$_{K_i}$, due to suppression of the background signal. An example for such case is the suppression of the non-specific $^{18}$F-FDG signal in the blood compartments allowing for a higher contrast in abnormal uptake regions, thereby confirming observations made in previous studies [24–27, 33]. This observation suggests increased likelihood of detecting metabolically active lesions surrounded by avid background non-specific uptake, such as in the case of intermediate differentiated HCCs with $^{18}$F-FDG imaging.

The Patlak graphical approach is of low linear complexity, thus allowing for fast computation and wide adoptability of WB parametric PET imaging in clinical practice. On the other hand, it assumes an irreversible $^{18}$F-FDG uptake, a condition not always satisfied in normal liver, kidneys, and certain tumors, such as HCCs [34], thereby leading to $K_i$ underestimation. Thus, an extended non-linear generalized Patlak model, considering a mild degree of reversibility, had been previously proposed [26] and could be easily integrated in our analysis. However, the use of a standard linear graphical analysis model is more preferable, thanks to its robustness to data noise, when the kinetic parameters are indirectly estimated from the dynamic PET images, and higher lesion contrast, when $K_i$ is underestimated only in the background.

We found a strong positive correlation ($\rho = 0.791$) between $K_{i\text{max}}$ and SUV$_{\text{max}}$, and even stronger correlations when looking separately at different organs. This suggests that $K_i$ images will identify hypermetabolic lesions if these are depicted on SUV images. In background neighboring regions, however, we found weak positive correlations ($\rho = 0.392$) between BKG$_{\text{mean}}$-$K_i$ and BKG$_{\text{mean}}$-SUV, which is expected owing to blood pool suppression within $K_i$ images.

The TBR and CNR are almost always higher in $K_i$ images (Fig. 5). On a purely visual examination, most of the malignant lesions were depicted on both modalities, except two HCC nodules appearing as non-hypermetabolic on both modalities and one HCC only depicted on $K_i$ images. Thus, in the majority of cases, the TBR$_{\text{SUV}}$ was largely sufficient to detect the lesions.

In the case of the lung lesions, for example, where the highest TBR values were found, we observe that most lung lesions are associated with a very low metabolism, resulting in lung lesions being easily detected. Overall, a TBR value between 3 and 5 is generally adequate in differentiating the lesion from its background, if the lesion has sufficient metabolic activity. Therefore, improving the TBR beyond this range with parametric imaging will not further improve lesion detectability, although it may improve quantification. On the other hand, detectability may be challenging for lesions with low TBR$_{\text{SUV}}$ scores especially within a highly vascularized organ of high BKG$_{\text{mean}}$ uptake, such as the liver. Detectability may then be improved by the higher TBR$_{K_i}$, due to suppression of the background signal. An example for such case is the

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Spearman’s $\rho$ and $p$ values (Sig. 2-tailed) between SUV and $K_i$ images, calculated for the maximum values of the lesions and the mean values of the backgrounds. Correlation is deemed significant for $p &lt; 0.01$. All values are reported for the full dataset and for lesions grouped by localization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spearman’s $\rho$ lesions</td>
</tr>
<tr>
<td>Total</td>
<td>0.777 ($p &lt; 0.001$)</td>
</tr>
<tr>
<td>Abdominal</td>
<td>0.796 ($p &lt; 0.001$)</td>
</tr>
<tr>
<td>Liver</td>
<td>0.890 ($p &lt; 0.001$)</td>
</tr>
<tr>
<td>Lungs</td>
<td>0.780 ($p = 0.001$)</td>
</tr>
<tr>
<td>Bones</td>
<td>0.950 ($p &lt; 0.001$)</td>
</tr>
<tr>
<td>Lymph nodes</td>
<td>0.657 ($p &lt; 0.001$)</td>
</tr>
<tr>
<td>Other</td>
<td>0.806 ($p = 0.005$)</td>
</tr>
</tbody>
</table>

Fig. 5 Whisker plots showing TBR and CNR for the SUV and $K_i$ images, presented by organ. $K_i$TBR values are always higher than SUV$_{\text{max}}$ values, regardless of the anatomical region, although the difference is variable depending on the organ. The same finding is observed for CNR

$\rho =$ Spearman’s correlation coefficient, $p =$ significance level.

<table>
<thead>
<tr>
<th>Modality</th>
<th>SUV</th>
<th>$K_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesion Localization</td>
<td>TBR</td>
<td>SUV</td>
</tr>
<tr>
<td>Lymph nodes</td>
<td>10000.00</td>
<td>10000.00</td>
</tr>
<tr>
<td>Lungs</td>
<td>10000.00</td>
<td>10000.00</td>
</tr>
<tr>
<td>Bones</td>
<td>10000.00</td>
<td>10000.00</td>
</tr>
<tr>
<td>Abdominal</td>
<td>10000.00</td>
<td>10000.00</td>
</tr>
<tr>
<td>Liver</td>
<td>10000.00</td>
<td>10000.00</td>
</tr>
<tr>
<td>Other</td>
<td>10000.00</td>
<td>10000.00</td>
</tr>
</tbody>
</table>

Eur Radiol
HCC (patient no. 1) appearing as non-metabolic in the SUV images and as hypermetabolic in the \( K_i \) images (Fig. 3).

However, our findings suggest that the TBR and CNR metrics may not always be sufficient for quantifying lesion detectability. However, they can both indicate ability to distinguish focal lesions from their background and draw appropriate regions of interest for further quantitative evaluations. In the case of the two hepatic lesions that were negative in both SUV and \( K_i \) images, the TBR\(_{SUV} \) were 0.41 and 0.32 whereas the respective TBR\(_{K_i} \) scores were 53.33 and 43.45, respectively. Moreover, the CNR\(_{SUV} \) values were 1.40 and 1.08 whereas the CNR\(_{K_i} \) values were 364.55 and 296.99, respectively. Despite the very high TBR and CNR scores of the \( K_i \) images, the lesions were not detected, due to the very low target lesion and nearly zero background \( K_i \) values. The same effect explains why the false-positive lung lesion was not detected on \( K_i \) images (Fig. 4).

There was one interesting case (patient no. 8) of biopsy-proven benign lung lesion which was erroneously depicted as positive in SUV images (Fig. 4). There were also several findings in the standard report of SUV images of muscle activity, inflammation, or benign lymph nodes, which were not visible on the \( K_i \) images. One example is the case of benign thoracic lymph nodes depicted on SUV images (patient no. 14) but not on \( K_i \) images (Supplementary Fig. 1). Overall, \( K_i \) images detected less false-positive and less benign lesions than SUV images, thereby improving specificity.

The most important limitation of this study is the relatively low number of patients. Nevertheless, we identified a large number of oncologic lesions with large histologic diversity to form a large variety of malignancies in order to drive useful conclusions for the clinical performance assessment of WB Patlak \( K_i \) \(^{18}\)F-FDG PET imaging. Moreover, having neglected the \( k_4 \) parameter during the Patlak analysis may have resulted in quantification errors, but only in the few regions, such as the liver or specific malignant tumors where non-negligible \(^{18}\)F-FDG uptake reversibility was expected. Furthermore, the manual ROI drawing for the IDIF calculation is user-dependent and can be subject to variability, although the net effect on the robust Patlak analysis was non-significant.

Although the proposed protocol requires approximately 50 min of patient’s involvement time, compared to 80 min for a standard SUV examination, the patient has to lie on the table during the whole 50-min examination. This may be challenging for patients and can impact throughput in clinical routine. We plan to continue this work with a prospective controlled study of a larger cohort targeting specific tumor types.

Acknowledgments This work was supported by the Swiss National Science Foundation under grant SNF3 320030_176052 and the Swiss Cancer Research Foundation under Grant KFS-3855-02-2016.

Funding This study has received funding by the Swiss National Science Foundation.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Prof. Habib Zaidi.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.
Ethical approval Institutional Review Board approval was obtained.

Methodology
- prospective
- observational
- performed at one institution

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References