ORIGINAL RESEARCH

Open Access

CrossMark

A new methodology to derive 3D kinetic parametric FDG PET images based on a mathematical approach integrating an error model of measurement

Elyse Colard[1], Sarkis Delcourt[2], Laetitia Padovani[3], Sébastien Thureau[4], Arthur Dumouchel[5], Pierrick Gouel[5], Justine Lequesne[6], Bardia Farman Ara[2], Pierre Vera[4], David Taïeb[7], Isabelle Gardin[8], Dominique Barbolosi[9†] and Sébastien Hapdey[8*†] (iD)

Abstract

Background: In FDG-PET, SUV images are hampered by large potential biases. Our aim was to develop an alternative method (ParaPET) to generate 3D kinetic parametric FDG-PET images easy to perform in clinical oncology.

Methods: The key points of our method are the use of a new error model of PET measurement extracted from a late dynamic PET acquisition of 15 min, centered over the lesion and an image-derived input function (IDIF). The 15-min acquisition is reconstructed to obtain five images of FDG mean activity concentration and images of its variance to model errors of PET measurement. Our approach is carried out on each voxel to derive 3D kinetic parameter images. ParaPET was evaluated and compared to Patlak analysis as a reference. Hunter and Barbolosi methods (Barbolosi-Bl: with blood samples or Barbolosi-Im: with IDIF) were also investigated and compared to Patlak. Our evaluation was carried on K_i index, the net influx rate and its maximum value in the lesion ($K_{i,max}$).

Results: This parameter was obtained from 41 non-small cell lung cancer lesions associated with 4 to 5 blood samples per patient, required for the Patlak analysis. Compare to Patlak, the median relative difference and associated range (median; [min;max]) in $K_{i,max}$ estimates were not statistically significant (Wilcoxon test) for ParaPET (− 3.0%; [− 31.9%; 47.3%]; $p = 0.08$) but statistically significant for Barbolosi-Bl (− 8.0%; [− 30.8%; 53.7%]; $p = 0.001$), Barbolosi-Im (− 7.9%; [− 38.4%; 30.6%]; $p = 0.007$) or Hunter (32.8%; [− 14.6%; 132.2%]; $p < 10^{-5}$). In the Bland-Altman plots, the ratios between the four methods and Patlak are not dependent of the K_i magnitude, except for Hunter. The 95% limits of agreement are comparable for ParaPET (34.7%), Barbolosi-Bl (30.1%) and Barbolosi-Im (30.8%), lower to Hunter (81.1%). In the 25 lesions imaged before and during the radio-chemotherapy, the decrease in the FDG uptake (ΔSUV_{max} or $\Delta K_{i,max}$) is statistically more important ($p < 0.02$, Wilcoxon one-tailed test) when estimated from the K_i images than from the SUV images (additional median variation of − 2.3% [− 52.6%; + 19.1%] for $\Delta K_{i,max}$ compared to ΔSUV_{max}).

(Continued on next page)

---

* Correspondence: sebastien.hapdey@chb.unicancer.fr
†Dominique Barbolosi and Sébastien Hapdey contributed equally to this work.
[8]Department of Nuclear Medicine, Centre Henri Becquerel, Rouen and LITIS-QuantIF-EA4108, University of Rouen, Rouen, France
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusion: None of the four methodologies is yet ready to replace the Patlak approach, and further improvements are still required. Nevertheless, ParaPET remains a promising approach, offering a non-invasive alternative to methods based on multiple blood samples and only requiring a late PET acquisition. It allows deriving K_i values, highly correlated and presenting the lowest relative bias with Patlak estimates, in comparison to the other methods we evaluated. Moreover, ParaPET gives access to quantitative information at the pixel level, which needs to be evaluated in the perspective of radiomic and tumour response.

Trial registration: NCT 02821936; May 2016.

Keywords: Dynamic PET, Parametric imaging, FDG kinetics, Lung cancer, Quantification

Introduction

Positron emission tomography using 18-fluorodeoxy-glucose (FDG PET) is useful for tumour staging, radiotherapy planning, treatment response and disease progression assessment [1].

FDG uptake is characterised by the standardised uptake value (SUV), which is widely used in clinical practice, based on a late static image acquired 60 min post-injection [2]. Therefore, different factors (both technical and physiological) can affect SUV calculation because it does not accurately represent the exact glucose metabolic rate, nor differentiate metabolised from unmetabolised FDG within the tumour [3]. Moreover, SUV is sensitive to patient preparation, scanning procedure, image reconstruction and image analysis procedures, compromising its use for pre-, per- and post-treatment scans comparison and inter-patient comparison [3, 4].

An alternative consists in the determination of the total behaviour of glucose within the lesion using more quantitative measurements. The gold standard for modelling tissue time-activity concentration curves derived from dynamic FDG PET acquisitions is the full kinetic analysis with compartmental modelling using a nonlinear least-squares regression [5] or the simplified Patlak graphical analysis [6]. These methods provide access to FDG kinetic parameters such as the net influx rate constant (K_i). Several studies have demonstrated the potential added value of kinetic parameters over SUV measurement for the assessment of disease progression [7]. Despite Patlak rapid calculation and simple expression, the method has two major practical limitations: the need for continuous blood sampling and a 1-h dynamic acquisition. Several derived Patlak analyses using image-derived input function (IDIF) extracted from the images of the aorta or the left ventricular to create the shape of the FDG blood concentration function $C_p(t)$ have been proposed in the literature [8–10].

As an alternative to Patlak approach, several simplified quantitative methods have been proposed [11, 12], such as Hunter method based on a static acquisition and a venous blood sample to scale the tri-exponential blood activity curve for each patient. Recently, Barbolosi et al.

[13] proposed a methodology to calculate global FDG kinetic parameters for the whole lesion, taking into account measurement errors. However, all these methods required blood samples.

In the literature, refined quantitative methods have been proposed to compute 3D kinetic parametric FDG PET images [13–15]. However, these methods required complex reconstruction algorithms and long PET acquisition that have prevented their clinical adoption.

Our aim was to develop a method to generate 3D kinetic parametric FDG PET images easy to perform in clinical routine. This method named ParaPET is based on previous work from Barbolosi et al. [13]. The main improvements of our method are the use of a new error model of PET measurement, an IDIF from a late dynamic acquisition and the measurement of kinetic parameters at the voxel level. This method is evaluated and compared to a derived Patlak analysis as a reference, and with Hunter and Barbolosi methods (Barbolosi-Bl: with blood samples or Barbolosi-Im: with IDIF).

Material and methods

Methods

To determine the FDG kinetic parameters within the lesion, the linear approximation of the mathematical representation of the standard three-compartment model with irreversible trapping ($k_4 = 0$) is considered according to Patlak analysis [6]. From $C_{FDG}(t_k)$, the FDG activity concentration in the lesion (Bq mL$_{\text{tissue}}^{-1}$) at a given time t_k after injection, the analytical solution of the three-compartment model verifies:

$$C_{FDG}(t_k) = K_i \int_0^{t_k} C_p(t)dt + V_p\, C_p(t_k) \qquad (1)$$

where $C_p(t_k)$ represents the FDG activity concentration in blood plasma at time t_k (Bq mL$_{\text{blood}}^{-1}$) and V_p the total blood distribution volume (i.e. the unmetabolized fraction of FDG in blood and interstitial volume, *unitless*). K_i, the net influx rate, is expressed as a combination of the compartmental transfer rates ($K_i = (k_1 \cdot k_3)/(k_2 + k_3)$). The K_i unit is mL$_{\text{blood}} \times$ mL$_{\text{tissue}}^{-1} \times$ min^{-1}.

Barbolosi approach

In this approach (further details in [13]), the K_i and V_p values are obtained by solving Eq. 1 for four time points t_k, considering that $C_{FDG}(t_k)$ and $C_p(t_k)$ are incorrect estimates of the true FDG activity concentration in the PET images and blood draws. The measurement variability of $C_{FDG}(t_k)$ and $C_p(t_k)$ is taken into account, leading to the following mathematical expressions:

$$C_{FDG}(t_k) = \overline{C_{FDG}(t_k)} + \varepsilon_{FDG,t_k} \qquad (2)$$

$$C_p(t_k) = \overline{C_p(t_k)} + \varepsilon_p \qquad (3)$$

where $\overline{C_{FDG}(t_k)}$ is the mean maximum FDG activity concentration in the lesion, averaged over three consecutive 2D slices (see Fig. 1) and ε_{FDG,t_k} the associated error, at study time t_k. $\overline{C_p(t_k)}$ is the mean FDG blood activity concentration measured from four blood samples, and ε_p the associated error derived from the counting efficiency which is considered identical for the same four time points t_k. The experimental errors ε_{FDG,t_k} and ε_p are distributed normally around 0 with σ^2_{FDG,t_k} and σ^2_p, their respective variance [13].

For each time t_k, 10,000 couples of ($C_{FDG}(t_k)$, $C_p(t_k)$) are computed from 10,000 random draws of ε_{FDG,t_k} and ε_p defined according to their respective distribution. Then, 10,000 couples of x and y values are obtained by minimising the following function:

$$f(x,y) = \sum_{i=1}^{4} \left(x \int_0^{t_k} C_p(\tau)d\tau + yC_p(t_k) - C_{FDG}(t_k) \right)^2 \qquad (4)$$

with $0 \leq x \leq K_H$ and $0 \leq y$, where K_H is derived from the

Hunter model [12]. It corresponds to the estimation of K_i index at the first study time point t_1 and is used as the maximal boundary in the search of x. A single couple of K_i and V_p values, corresponding to the mean values of x and y, respectively, is used to characterise the whole lesion.

ParaPET approach

Our proposed approach is a non-invasive extension of Barbolosi method at the voxel level, integrating a new error model of measurement, yielding to the determination of 3D parametric images of the lesion.

To improve the error model of measurements, we propose to compute five images of FDG mean activity concentration, associated with five images of its variability from FDG PET images centered over the lesion.

To that end, a delayed 15 min FDG dynamic PET acquisition performed at 80–90 min post-injection and centered over the lesion is necessary. This acquisition is resampled into five series of 3 min listmode datasets. For each 3 min listmode dataset, five reconstructions of 2 min are generated. For example, for the first 3 min listmode bin (t_1), we reconstructed the data from 0 to 2 min, 15 s–2 min 15 s, 30 s–2 min 30 s, 45 s–2 min 45 s, and 1–3 min. From our 15 min listmode datasets, we finally generated 25 reconstructed series. This allows the determination of $\overline{C_{FDG}(t_k)}$ and σ^2_{FDG,t_k} ($k = 1$ to 5) variables at the voxel level needed for Eq. 2 (cf Fig. 2).

Since we propose a methodology applicable at the voxel level, it is mandatory to avoid any misalignment between the 25 reconstructions used to derive $\overline{C_{FDG}(t_k)}$ and σ^2_{FDG,t_k}. A precise registration of the series must be made beforehand. In order to minimise the offset
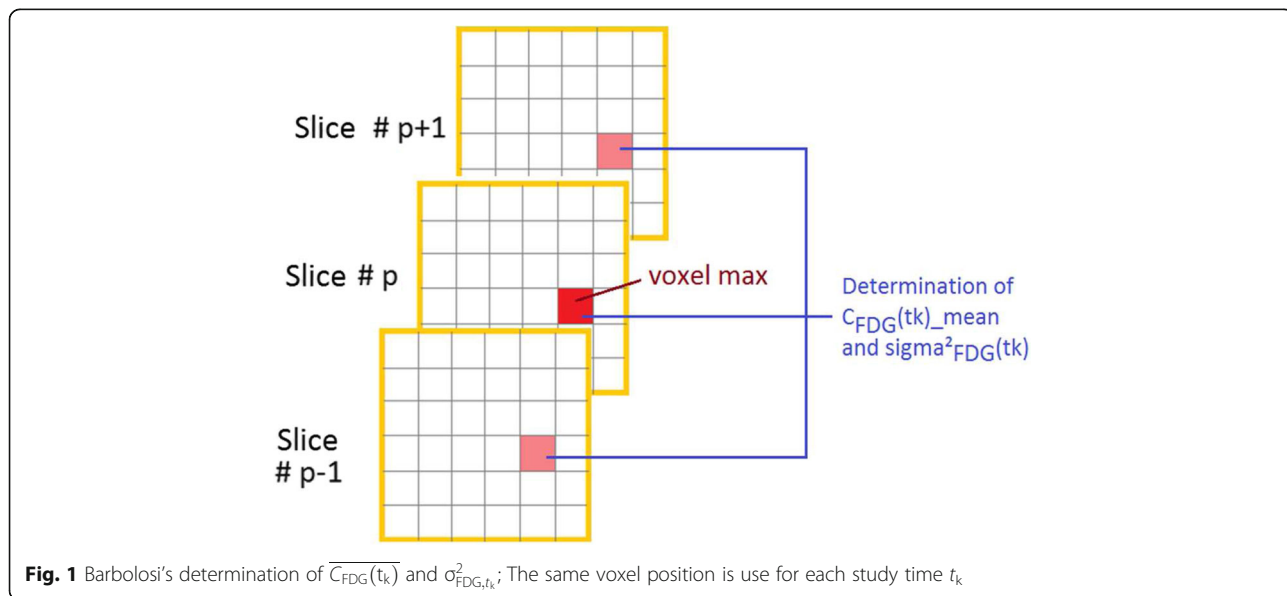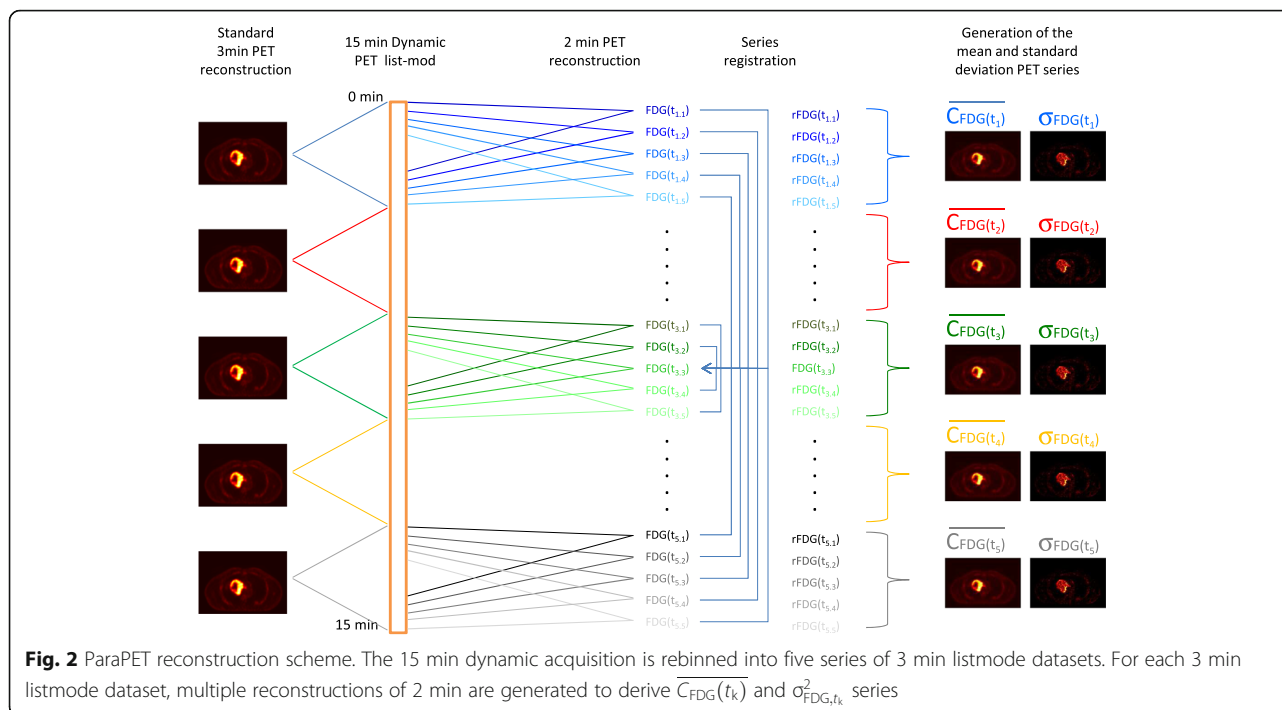

Fig. 1 Barbolosi's determination of $\overline{C_{FDG}(t_k)}$ and σ^2_{FDG,t_k}; The same voxel position is use for each study time t_k

Fig. 2 ParaPET reconstruction scheme. The 15 min dynamic acquisition is rebinned into five series of 3 min listmode datasets. For each 3 min listmode dataset, multiple reconstructions of 2 min are generated to derive $\overline{C_{FDG}(t_k)}$ and σ^2_{FDG,t_k} series

applied to each series, all data were locally realineated on the central time series ($k = 3$) using a 3D rigid registration method based on mutual information.

To avoid multiple blood samples, a late IDIF from FDG PET images is determined. Five regions of interest (ROI) are manually drawn over the aorta at each study time t_k, allowing the extraction of the mean activity concentration $\overline{C_p(t_k)}$ and its variance σ^2_{p,t_k} needed for the error model in Eq. 3. The shape of the $C_p(t)$ function is modelled using the tri-exponential function developed in Hunter's method.

Finally, the equation resolution following the approach previously proposed by Barbolosi (Eq. 4) is carried out on each voxel to derive parametric PET images of FDG kinetic parameters.

In practice, to perform our non-invasive ParaPET approach, only a late dynamic PET acquisition of 15 min performed at 80–90 min post-injection and centered over the lesion is necessary.
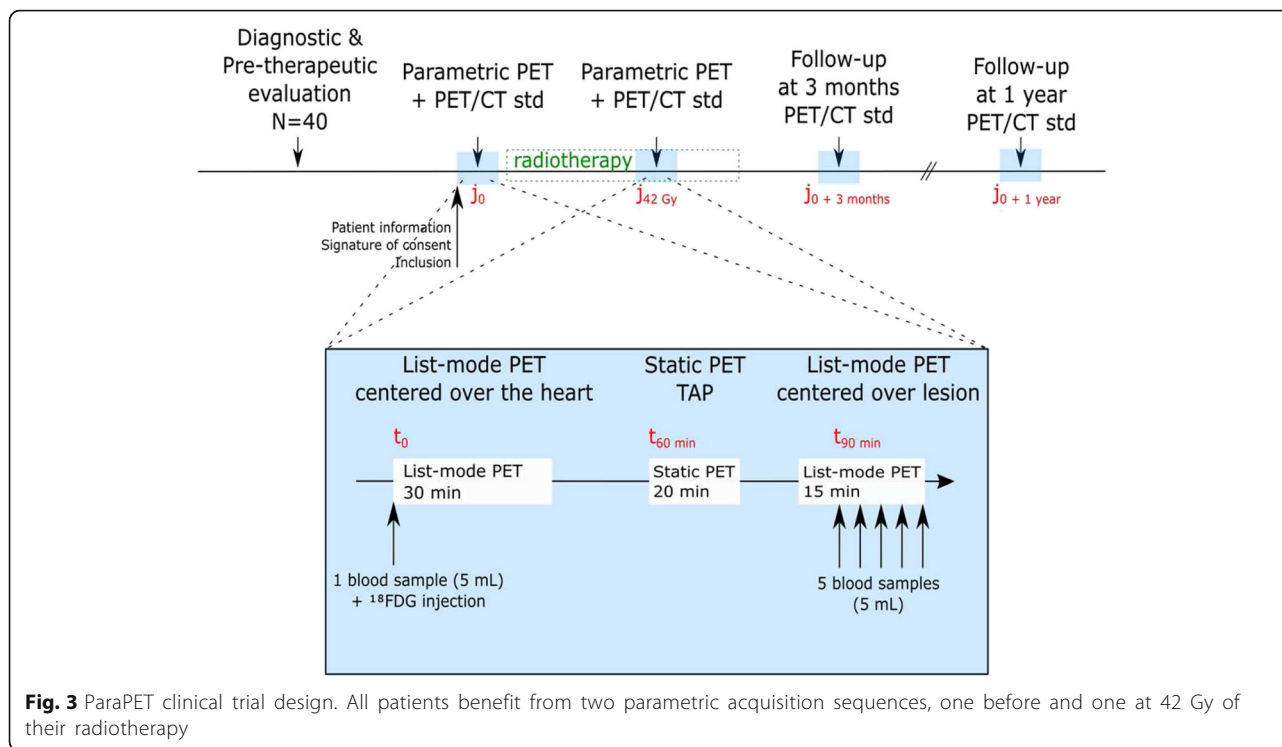
Method evaluation

Our method was evaluated on patients suffering for stage III non-small cell lung cancer (NSCLC) included in the ParaPET clinical trial (NCT 02821936). The study protocol was approved by the local Research Ethics Committee and has therefore been performed in accordance with the ethical standards laid down on the 1964 Declaration of Helsinki and its later amendments. All patients gave their written informed consent. Each

patient benefits from 2 PET-CT exams, the baseline exam performed before any treatment and a per-therapeutic exam, around 42 Gy (21 days) of the radiotherapy, as indicated in Fig. 3.

All patients benefit from two parametric acquisition sequences, performed on a Discovery 710 PET/CT device (GE Healthcare, Milwaukee, WI, USA). The acquisition sequence is divided into three acquisitions (cf. Fig. 3):

1) An early 30 min listmode PET acquisition centered over the heart, starting simultaneously with FDG injection of 3.5 MBq/kg. This acquisition which is required for the Patlak analysis only, was reconstructed using a 30 frames protocol [16];
2) A 20-min standard whole-body PET acquisition was carried out 60 min post injection. These clinical routine images are not used in our protocol;
3) Lastly, a 15 min listmode PET acquisition centered over the lesion (performed 80–90 min post-injection). Five late venous blood samples are required for the implementation of the approaches of Hunter, Barbolosi and Patlak. The blood withdraws were realised each 3 min, starting 1.5 min after the beginning of the listmode acquisition.

Our evaluation was carried on K_i index, the net influx rate and its maximum value in the lesion ($K_{i,max}$). We evaluated the robustness of our method with regards to the number of random draws of ε_{FDG,t_k} and ε_{p,t_k} by

Fig. 3 ParaPET clinical trial design. All patients benefit from two parametric acquisition sequences, one before and one at 42 Gy of their radiotherapy

repeating calculation 5 times using 10 to 10,000 random draws. To this end, we derived the standard deviation associated with the 5 repeated calculation of $K_{i,max}$.

From the optimal number of random draws, $K_{i,max}$ value measured by our method was compared to the reference K_i value derived from Patlak analysis considering the voxel with the maximal intensity. The same comparison was done with Hunter and Barbolosi methods (Barbolosi-Bl or Barbolosi-Im).

To compute reference K_i, using the Patlak approach, the early 30 min PET acquisition centered over the heart is required, along with the late 15 min PET acquisition centered over the lesion acquired 80–90 min post-injection and 5 venous blood samples (each 3 min).

Hunter and Barbolosi original methods (Barbolosi-Bl) required the late 15 min acquisition reconstructed as 5×3 min acquisitions associated with venous blood samples. Whole blood activity was measured using a high energy dedicated Gamma Counter.

Barbolosi-Im and ParaPET approaches only require a late dynamic PET acquisition centered over the lesion without any blood sampling.

To assess the agreement between the Patlak K_i and $K_{i,max}$ measured by all methods, Bland-Altman ratios ($K_{i,max}$/Patlak K_i as a function of averaged value of $K_{i,max}$ and Patlak K_i) and intraclass correlation coefficients (ICC) were calculated. The statistical comparison of the methods was assessed by a Wilcoxon paired test. The 95% limits of agreement on the

Bland-Altman plots were expressed as a percentage of the mean K_i ratios.

Considering the whole set of lesions, $K_{i,max}$ obtained with ParaPET was compared to the lesion SUV_{max} index measured on a 3 min acquisition at time $t_{k=1}$. The agreement was evaluated using Pearson correlation coefficient and its statistical significance by the p value.

We also investigated the influence of $K_{i,max}$ and SUV_{max} in the tumour response evaluation, considering the $\Delta K_{i,max}$ and ΔSUV_{max} (unitless) between the baseline and per-therapeutic exams:

$$\Delta K_{i,\,max} = \left(K_{i,42Gy} - K_{i,baseline}\right)/Ki, baseline$$

$$\Delta SUV_{max} = \left(SUV_{max,42Gy} - SUV_{max,baseline}\right)/SUV_{max,baseline}$$

Bland-Altman plots and ICC were calculated. The statistical comparison was assessed by a Wilcoxon signed-rank test.

Results

Table 1 gives the patient demographics. To date, 31 patients were included in this on-going clinical trial, yielding a total of 52 parametric PET sequences (31 baseline and 18 per-therapeutic). The total number of lesions studied is 82 (51 in baseline, 31 in per-therapeutic exams).

Among the 52 exams realised, we were able to obtain at least four blood samples for only 52% of our exams. This result clearly shows how difficult it is to implement

Table 1 Patient demographics and characteristics

Number of patients		31
Number of exams		52
Male/female (%)		72 /28
Age (years)		63.1 ± 8.3
Histology	ADK	42%
	SCC	48%
	LCLC	10%
Location	Right	
	Upper	39%
	Middle	13%
	Lower	10%
	Left	
	Upper	29%
	Lower	10%
Average failure rate of blood sampling	No samples or unusable	35%
	1 to 3 samples	13%
	≥ 4 samples	52%

methods requiring blood samples. Therefore, we were finally only able to compare the ParaPET methodology with Patlak, Hunter and Barbolosi for 41/82 lesions.

Regarding the comparison between ParaPET images and standard PET images, 82 lesions were included and 25 lesions with a complete baseline and per-therapeutic 42 Gy sequences.

Fig. 4a, b shows the results of the influence of the number of random draws on the robustness of the estimation of $K_{i,max}$ and $V_{p,max}$ values for the eight first lesions studied. The same behaviour was found with the other lesions. The method clearly presents a fast convergence t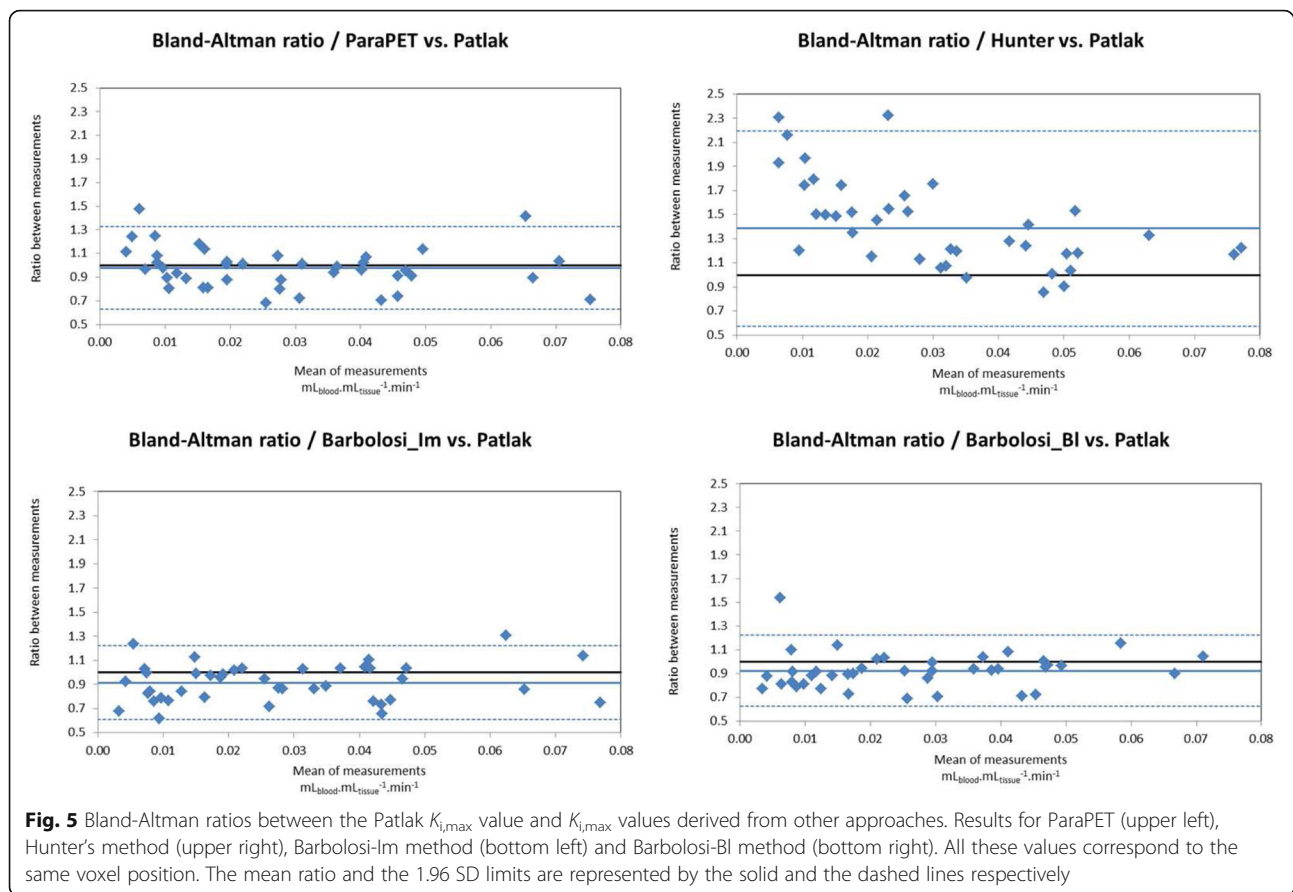o a stable solution for K_i and V_p. Considering a maximum coefficient of variation (CV) of 3%, our results show that 1000 random draws are required to correctly estimate $K_{i,max}$ and V_p.

Regarding the estimation of the $K_{i,max}$ index, all methods show a good correlation with Patlak (ICC of 0.905 for Hunter and > 0.935 for the other approaches). Compare to Patlak, the median relative difference and associated range (median; [min;max]) in $K_{i,max}$ estimates were not statistically significant (Wilcoxon test) for Para-PET (− 3.0%; [− 31.9%;47,3%]; $p = 0.08$) but statistically significant for Barbolosi-Bl (− 8.0%; [− 30.8%;53.7%]; $p = 0.001$), Barbolosi-Im (− 7.9%; [− 38.4%;30.6%]; $p = 0.007$) or Hunter (32.8%; [− 14.6%;132.2%]; $p < 10{-}5$). Figure 5 displays the Bland-Altman ratios between the Patlak value of $K_{i,max}$ and the values found with ParaPET, Barbolosi-Bl and Barbolosi-Im. This figure shows more clearly the large bias observed with the Hunter's methodology, and its dependence with the K_i amplitude that was not observed for Barbolosi or ParaPET methods. The 95% limits of agreement are comparable for ParaPET (34.7%), Barbolosi-Bl (30.1%) and Barbolosi-Im (30.8%), but much lower than Hunter's (81.1%) and do not highlight the superiority of Para-PET nor Barbolosi's approaches.

Figs. 6 and 7 show examples of SUV and K_i PET images generated by our approach. For comparison purpose, K_i and SUV images were normalised to the total number of counts measured in the slice for Fig. 6, or in the background activity of the SUV image for Fig. 7. These images are highly correlated as it was previously reported [17], even if small differences can be observed at the voxel level as it can be seen in the subtraction image (see Fig. 6).

As suggested by other groups, the K_i images present a higher image contrast, as it can be seen on the profiles drawn through the lesion on Fig. 7,
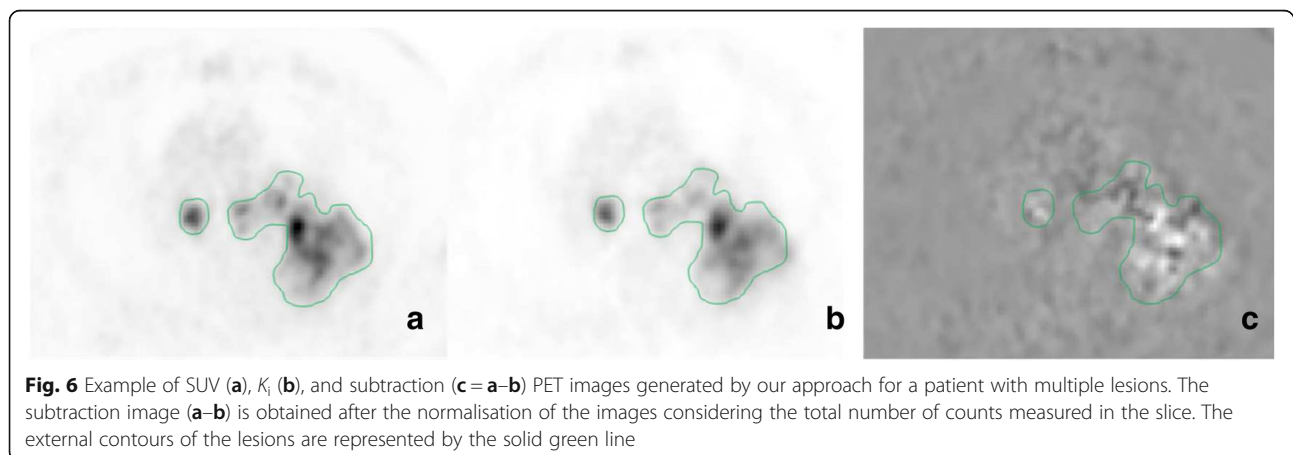

Fig. 4 Evaluation of the robustness of our proposed method; $K_{i,max}$ (left fig.) and $V_{p,max}$ (right fig.) estimates using 10 to 10,000 random draws for 8 lesions. Each calculation was repeated five times to derive the standard deviation on $K_{i,max}$ estimates. The coefficients of variation of 1% and 3% are represented by the dotted and the dashed lines respectively

Fig. 5 Bland-Altman ratios between the Patlak $K_{i,max}$ value and $K_{i,max}$ values derived from other approaches. Results for ParaPET (upper left), Hunter's method (upper right), Barbolosi-Im method (bottom left) and Barbolosi-Bl method (bottom right). All these values correspond to the same voxel position. The mean ratio and the 1.96 SD limits are represented by the solid and the dashed lines respectively

especially within the necrotic region of the lesion. In this example, the contrast between lesion maximum uptake and the necrotic region was improved from 5.1 in the SUV image to 5.7 in the K_i image.

The correlation plot between the SUV_{max} and $K_{i,max}$ index for the 52 examinations is displayed on Fig. 8. This figure clearly highlights the strong correlation between the 2 indexes.

In complement, Fig. 9 shows the difference between $\Delta K_{i,max}$ and ΔSUV_{max} index which represent the variation of FDG uptake before (J0) and during radiochemotherapy (J42 Gy). In the 25 lesions imaged before and during the radio-chemotherapy, the $\Delta K_{i,max}$ is statistically lower than the ΔSUV_{max} ($p < 0.02$ Wilcoxon one-tailed test). The linear relation between those indexes is given by $\Delta K_{i,max} = 0.84 \ \Delta SUVmax-$

Fig. 6 Example of SUV (**a**), K_i (**b**), and subtraction (**c** = **a**–**b**) PET images generated by our approach for a patient with multiple lesions. The subtraction image (**a**–**b**) is obtained after the normalisation of the images considering the total number of counts measured in the slice. The external contours of the lesions are represented by the solid green line
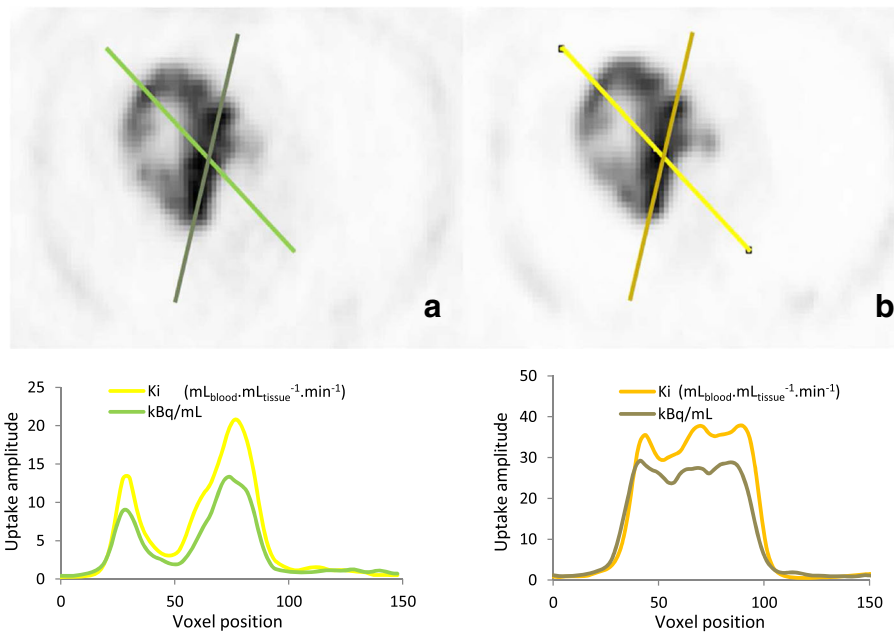
Fig. 7 Example of profiles drawn through the lesion of the SUV (**a**) and K_i (**b**) PET images; K_i and SUV images were normalised to the background activity of the SUV image
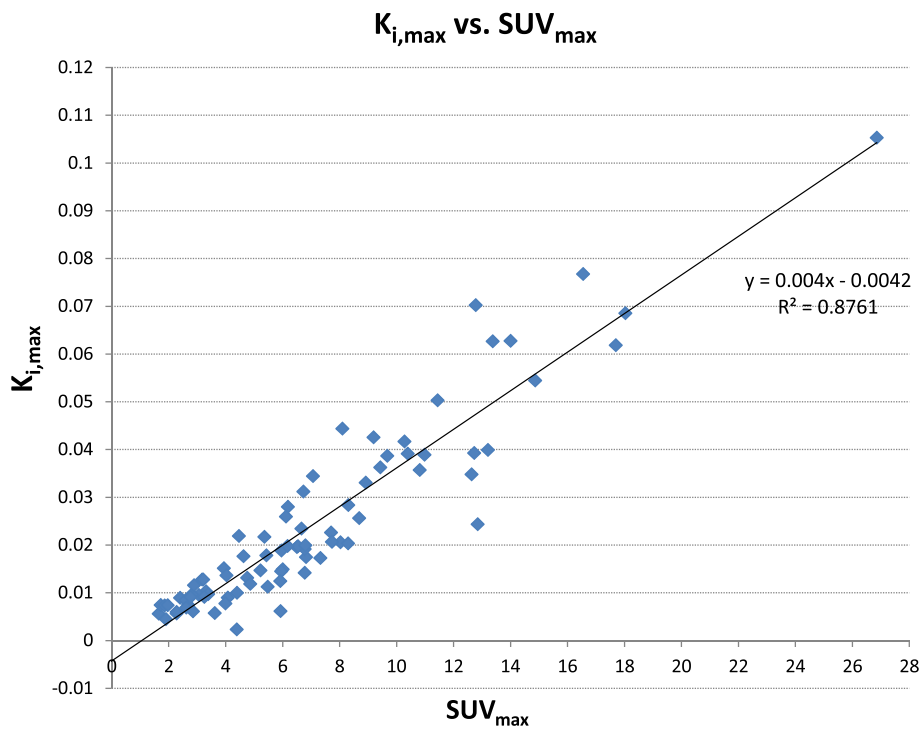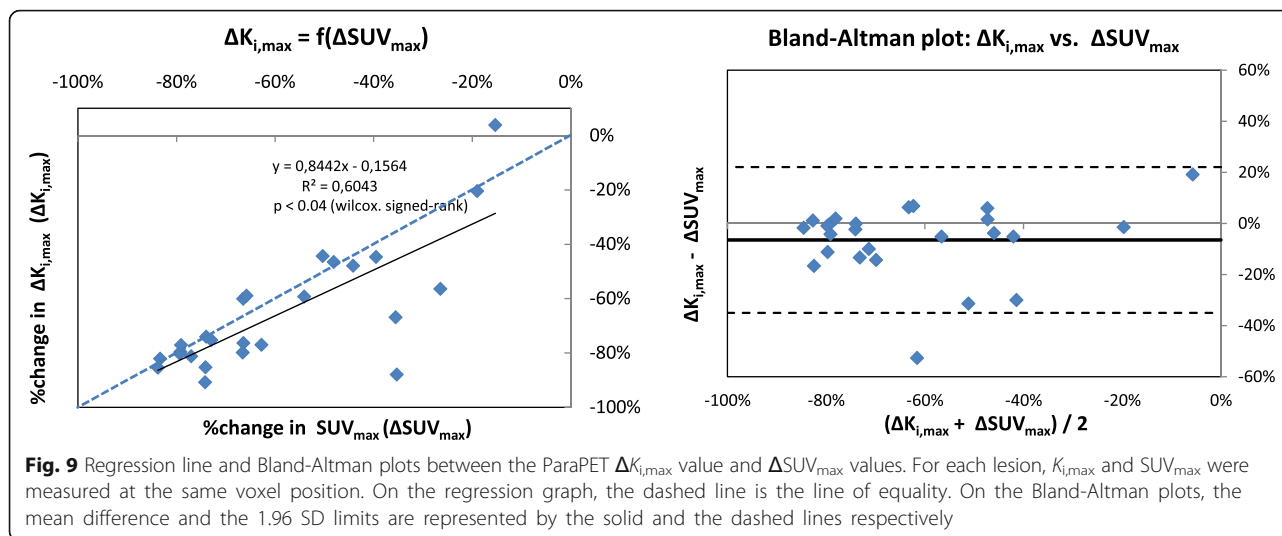

Fig. 8 Regression line between the ParaPET $K_{i,max}$ value and SUV$_{max}$ values

Fig. 9 Regression line and Bland-Altman plots between the ParaPET $\Delta K_{i,max}$ value and ΔSUV_{max} values. For each lesion, $K_{i,max}$ and SUV_{max} were measured at the same voxel position. On the regression graph, the dashed line is the line of equality. On the Bland-Altman plots, the mean difference and the 1.96 SD limits are represented by the solid and the dashed lines respectively

0.15 with a median additional variation of -2.3% [-52.6%; $+19.1\%$].

Discussion

This paper proposes a new non-invasive methodology, called ParaPET, allowing the determination of 3D map of FDG kinetic parameters. Our approach is adapted to clinical routine and requires only a late dynamic PET acquisition of a limited duration corresponding to 15 min, performed after the whole-body PET/CT acquisition, 80 min after FDG injection in our evaluation. Furthermore, blood samples are not necessary to perform our method.

In FDG PET, kinetic parameters are estimated using the Patlak graphical analysis as a reference, according to the European Organization for Research and Treatment of Cancer (EORTC) [18]. Although Patlak analysis is robust, the original method has three major practical limitations: the need for continuous blood sampling, a 1-h dynamic acquisition and the assumption of irreversible trapping. Several alternatives have been proposed, but they still require a long dynamic PET acquisition and the inclusion of blood compartments in the field of view [19, 20]. In this study, the implementation of Patlak analysis is performed using a shortened dynamic FDG PET protocol based on an early 30 min dynamic PET acquisition starting from the injection of FDG and a late 15 min PET sequence acquired 80 min post injection. The shape of the FDG blood activity concentration function $C_p(t)$ is derived from the early dynamic PET acquisition and five late blood measurements.

In the literature, refined quantitative methods based on the Markov chain model [21] or direct 4D parametric PET images using expectation-maximisation [15] or penalised maximum-likelihood [14] reconstruction algorithms have been proposed allowing accurate calculation

of kinetic parameters. Although these kinetic modelling methods are powerful, their complexity, the slow convergence rate of reconstruction algorithms and the need of long PET acquisition have prevented their clinical adoption.

Recently, as an alternative to Patlak approach, Barbolosi et al. [13] proposed a methodology for the determination of kinetic parameters such as the metabolised and unmetabolised FDG fractions. This method, evaluated on paragangliomas [13] and on NSCLC [22], takes into account the variability of activity concentration measurements. However, three major limitations have to be highlighted. First, the method was designed to compute global kinetic parameter values for the whole lesion, whereas parametric images at a voxel level might provide additional quantitative information. Second, the determination of $\overline{C_{FDG}(t_k)}$ and its variance σ^2_{FDG,t_k} is questionable, since values are derived from the maximum FDG activity concentration within three consecutive 2D slices surrounding the maximum activity concentration of the whole lesion (see Fig. 1). Third, this approach is invasive requiring several blood samples. We have faced these drawbacks by proposing a non-invasive methodology allowing the determination of FDG kinetic parameters at the voxel level, based on a new error model of measurement and a late IDIF.

As an alternative to blood samples, FDG activity concentration in blood can be extracted from dynamic PET images [15]. Regarding the determination of $\overline{C_{FDG}(t_k)}$ and its variance σ^2_{FDG,t_k}, our approach assumes that the FDG burden remain stable during each 3 min listmode datasets used to derived the mean FDG uptake and its associated variance. Our approach is naturally limited to lesions presenting an irreversible trapping since it is based on the Patlak kinetic model.

The validation of our methodology was carried out by taking the Patlak method as a reference rather than the full compartmental analysis. Unfortunately, the health status of the patients included in this protocol was not compatible with the realisation of 60 min dynamic acquisition required for the full compartmental approach. This also explains the low blood collection rate that we have been able to obtain in patients with severe reduction of their venous capital, especially after few weeks of radio-chemotherapy.

In the implementation of Barbolosi's, Hunter's and Patlak approaches, whole blood was used instead of plasma activity concentrations. This simplification was evaluated on the first four patients (four baseline and three per-therapeutic exams), comparing the measurements of the FDG concentration in the total blood or in the plasma (extracted by a 3000 rotation/min during 10 min). Seventy samples were compared (7 exams × 5 samples divided into 2 sub-samples). The relative difference was considered negligible (mean relative difference and associated standard deviation of $0.60 \pm 1.98\%$) and whole blood was finally used in the rest of the study.

The evaluation of the number of random draws showed that 1000 random draws are necessary to compute $K_{i,max}$ with a CV below 1% and $V_{p,max}$ with a CV below 3%.

In comparison to Hunter and Barbolosi methods, we showed that our approach presents the lowest relative bias in the estimation of the reference $K_{i,max}$. However, in the Bland-Altman ratio, the 95% limits of agreement remain too elevated (LoAs > 30%) to conclude in the replacement of the Patlak approach by any of the four methods. Nevertheless, even if our method does not seem to differ from Barbolosi's methods in terms of median error and range, it allows avoiding blood samples and requires only a late acquisition, making its clinical use possible. It also provides a map of K_i values, which is not possible with the other methods evaluated in this work.

The Bland-Altman is a graphical representation of the agreement between two methods or evaluators. It does not give any *p* value to assess the statistical significance of this agreement. We then decided to add the evaluation of the statistical difference between Patlak K_i values and the K_i estimates from the other approaches using a Wilcoxon paired test.

We assumed that the noise characteristic of our reconstructed image is correctly estimated using a simple resampling of our listmode data. The best approach would have been to derive the standard deviation of tumour absorption using a conventional bootstrap approach, generated by randomly drawing, with replacement, prompt and random events [23]. We then consider our solution as a particular case of a bootstrap approach.

Implementing a real bootstrap approach was not possible for this study but will be integrated as an evolution of our methodology.

The large variability of the errors in the estimation of K_i with ParaPET or Barbolosi's methods might be due to the relatively short time used to derive the K_i values (only 15 min). The inclusion of the whole-body acquisition could help to reduce the error and probably increase the accuracy of our methodology in comparison with Patlak, computed with the 0–30 min and the 80–95 min post-injection datasets.

Hunter's method presents the highest mean error in the estimation of K_i. This method was developed to be applied at 55 min post-injection. Since the blood activity is reduced at later time (80 min PI) and given that Hunter's method neglects the unmetabolized part of the FDG (i.e. $V_p = 0$), applying Hunter's method at later time may have a negative impact on its K_i estimates. Whatever the method, using late acquisitions with very low blood concentrations is very challenging since we derive the actual blood activity concentration, and the measure variability from those late series. However, due to practical consideration, using sooner series was not possible for this study. Nevertheless, the results even not perfect are encouraging and further improvements are still required. Here again, the inclusion of the whole-body acquisition would help to improve our approach.

As it was underlined in other papers [7, 24, 25], the K_i index, even highly correlated with the SUV values, gives additional information. When looking at the tumour response with therapy, tumour response based on changes in $K_{i,max}$ may thus be different than those measured with SUV_{max} values. The variation in $K_{i,max}$ is statistically higher than the variation in SUV_{max} values. Considering the EORTC or PERCIST criteria, the tumour response could be modified considering the $\Delta K_{i,max}$ instead of the ΔSUV_{max} value. For clinical trials, this could imply that use of SUV for assessment of radio-chemotherapy may under- or overestimate treatment efficiency compared with parametric analysis.

In addition, as shown in Fig. 6, the K_i and SUV images are not completely identical. The use of parametric images (K_i, V_p) for the determination of predictive and/or prognostic values based on radiomic features will remain to be investigated.

The main limitations of this study is the restriction of the validation of our 3D approach to $K_{i,max}$ corresponding to the maximum value of the net influx rate within the entire lesion instead of a full 3D evaluation.

Our approach is adapted to clinical routine since it only requires a dynamic PET acquisition of a limited duration corresponding to 15 min, performed 80 min after FDG injection in our evaluation. To date, our approach is limited to one bed position. In the next few

years, the new cristal-SiPM PET system, with larger field of view and increased sensitivity, will probably allow to implement the methodology for larger field of view.

Another limitation of our approach, to be applied in clinical routine, is the time needed to obtain parametric 3D maps. Efforts could be made to optimise the calculation time by using a parallelized process.

A PET system manufacturer recently proposes an automatic non-invasive Patlak whole-body acquisition in their scanner. Their approach even if it requires a 60-min acquisition, brings Patlak analysis back to the clinical applications. In that context, our approach could be a practical alternative to derive Patlak images only requiring a 15-min PET acquisition.

In a further step, the clinical relevance of our approach, already suggested in the paper from Padovani et al. [22] will be evaluated with the NSCLC patients included into the on-going ParaPET clinical trial (NCT 02821936). The aim of this trial is the assessment of the predictive and prognostic value of biomarkers extracted from the 3D map of kinetic parameters in the progression-free and overall survival. The results presented here have allowed us to validate our approach for the evaluation of prognostic and predictive values of parametric images.

Another major perspective of our approach is the evaluation of response to immunotherapy treatment. Indeed, so far, conventional PET imaging has shown its limits in the detection and interpretation of the tumour pseudo-progression due to an inflammatory process. Parametric imaging offers new insights into the use of FDG-PET for this purpose.

Conclusion

The main contribution of this study is the development of a non-invasive voxel-based approach for the determination of 3D parametric images of kinetic parameters of FDG uptake, integrating a new error model of PET measurement and a late IDIF. Our approach is not yet ready to replace the Patlak approach, and further improvements are still required. Nevertheless, ParaPET remains a promising approach, offering a non-invasive alternative to methods based on multiple blood samples and only requiring a simple 15 min listmode acquisition. It provides a viable alternative for the voxel-based estimation of the net influx rate K_i, giving access to additional heterogeneity information which could be used for radiomic purpose.

We confirmed in a population of patient treated by radio-chemotherapy that our ParaPET approach yields K_i indexes non-statistically different than Patlak K_i. As previously described, changes in SUV_{max} during the treatment is not correlated with the changes measured from K_i indexes.

Author details
[1]LITIS-QuantIF-EA4108, University of Rouen, Rouen, France. [2]Department of Nuclear Medicine, La Timone University Hospital, Marseille, France. [3]Department of Radiotherapy, La Timone University Hospital, Marseille and SMARTc-INSERM-UMR 911 CR02, Aix-Marseille University, Marseille, France. [4]Department of Radiotherapy, Centre Henri Becquerel, Rouen and LITIS-QuantIF-EA4108, University of Rouen, Rouen, France. [5]Department of Nuclear Medicine, Centre Henri Becquerel, Rouen, France. [6]Department of Clinical Research, Centre Henri Becquerel, Rouen, France. [7]Department of Nuclear Medicine, La Timone University Hospital, Marseille and CERIMED, Aix-Marseille University, Marseille, France. [8]Department of Nuclear Medicine, Centre Henri Becquerel, Rouen and LITIS-QuantIF-EA4108, University of Rouen, Rouen, France. [9]SMARTc–CRCM, INSERM UMR1068, CNRS UMR7258, Aix Marseille Université U105, Institut Paoli Calmette et APHM, Marseille, France.

References
1. Edet-Sanson A, Dubray B, Doyeux K, et al. Serial assessment of FDG-PET FDG uptake and functional volume during radiotherapy (RT) in patients with non-small cell lung cancer (NSCLC). Radiother Oncol. 2012;102:251–7.

2.   Boellaard R. Standards for PET image acquisition and quantitative data analysis. J Nucl Med. 2009;50(Suppl 1):11S–20S.

3.   Keyes JWJ. SUV: standard uptake or silly useless value? J Nucl Med. 1995;36:1836–9.

4.   Weber WA, Gatsonis CA, Mozley PD, et al. Repeatability of 18F-FDG PET/CT in advanced non-small cell lung cancer: prospective assessment in two multicenter trials. J Nucl Med. 2015;56(8):1137–43.

5.   Sokoloff L, Reivich M, Kennedy C, et al. The [14C]Deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure and normal values in the consious and anesthetized albino rat. J Neurochem. 1977;28:897–916.

6.   Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. J Cereb Blood Flow Metab. 1983;3:1–7.

7.   Cheebsumon P, Velasquez LM, Hoekstra CJ, et al. Measuring response to therapy using FDG PET: semi-quantitative and full kinetic analysis. Eur J Nucl Med Mol Imaging. 2011;38:832–42.

8.   de Geus-Oei LF, Visser EP, Krabbe PF, et al. Comparison of image-derived and arterial input functions for estimating the rate of glucose metabolism in therapy-monitoring 18F-FDG PET studies. J Nucl Med. 2006;47:945–9.

9.   Gambhir SS, Schwaiger M, Huang SC, et al. Simple noninvasive quantification method for measuring myocardial glucose utilization in humans employing positron emission tomography and fluorine-18 deoxyglucose. J Nucl Med. 1989;30:359–66.

10. Ohtake T, Kosaka N, Watanabe T, et al. Noninvasive method to obtain input function for measuring tissue glucose utilization of thoracic and abdominal organs. J Nucl Med. 1991;32:1432–8.

11. Hapdey S, Buvat I, Carson JM, Carrasquillo JA, Whatley M, Bacharach SL. Searching for alternatives to full kinetic analysis in 18F-FDG PET: an extension of the simplified kinetic analysis method. J Nucl Med. 2011;52:634–41.

12. Hunter GJ, Hamberg LM, Alpert NM, Choi NC, Fischman AJ. Simplified measurment of deoxyglucose utilization rate. J Nucl Med. 1996;37:950–5.

13. Barbolosi D, Hapdey S, Battini S, et al. Determination of the unmetabolised F-FDG fraction by using an extension of simplified kinetic analysis method: clinical evaluation in paragangliomas. Med Biol Eng Comput. 2015;54(1):103–11.

14. Kamasak ME, Bouman CA, Morris ED, Sauer K. Direct reconstruction of kinetic parameter images from dynamic PET data. IEEE Trans Med Imaging. 2005;24:636–50.

15. Karakatsanis NA, Casey ME, Lodge MA, Rahmim A, Zaidi H. Whole-body direct 4D parametric PET imaging employing nested generalized Patlak expectation-maximization reconstruction. Phys Med Biol. 2016;61:5456–85.

16. Hoekstra CJ, Hoekstra OS, Stroobants SG, et al. Methods to monitor response to chemotherapy in non-small cell lung cancer with 18F-FDG PET. J Nucl Med. 2002;43:1304–9.

17. Tixier F, Hatt M, Le Rest CC, le PA CL, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. J Nucl Med. 2012;53:693–700.

18. Young H, Baum R, Cremerius U, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. Eur J Cancer. 1999;35:1773–82.

19. Strauss LG, Pan L, Cheng C, Haberkorn U, Dimitrakopoulou-Strauss A. Shortened acquisition protocols for the quantitative assessment of the 2-tissue-compartment model using dynamic PET/CT 18F-FDG studies. J Nucl Med. 2011;52:379–85.

20. Visser EP, Kienhorst LBE, de Geus-Oei LF, Oyen WJG. Shortened dynamic FDG-PET protocol to determine the glucose metabolic rate in non-small cell lung carcinoma. IEEE Nucl Sci Symp Conf Rec. 2008:4455–8.

21. Huang J, O'Sullivan F. An analysis of whole body tracer kinetics in dynamic PET studies with application to image-based blood input function extraction. IEEE Trans Med Imaging. 2014;33:1093–108.

22. Padovani L, Baret A, Ciccolini J, et al. An alternative parameter for early forecasting clinical response in NSCLC patients during radiotherapy: proof of concept study. Br J Radiol. 2016;89:20160061.

23. Buvat I. A non-parametric bootstrap approach for analysing the statistical properties of SPECT and PET images. Phys Med Biol. 2002;47:1761–75.

24. Freedman N, Sundaram SK, Kurdziel K, et al. Comparison of SUV and Patlak slope for monitoring of cancer therapy using serial PET scans. Eur J Nucl Med. 2003;30:46–53.

25. Kotasidis FA, Tsoumpas C, Rahmim A. Advanced kinetic modelling strategies: towards adoption in clinical PET imaging. Clin Transl Imaging. 2014;2:219–37.