Spectral Clustering predicts tumor tissue heterogeneity using dynamic
18F-FDG PET: a complement to the standard compartmental modeling
approach

Prateek Katiyar1, 2, Mathew R. Divine1, Ursula Kohlhofer3, Leticia Quintanilla-Martinez3,
Bernhard Schölkopf2, Bernd J. Pichler1, and Jonathan A. Disselhorst1

1: Werner Siemens Imaging Center, Department of Preclinical Imaging and Radiopharmacy,
Eberhard Karls University Tuebingen, Tuebingen, Germany

2: Max Planck Institute for Intelligent Systems, Tuebingen, Germany

3: Institute of Pathology and Neuropathology, Eberhard Karls University Tuebingen and
Comprehensive Cancer Center, University Hospital Tuebingen, Tuebingen, Germany

Short running title: Clustering of dynamic PET tumor images

Word Count: 5045
**Corresponding author:**

Jonathan A. Disselhorst

University of Tuebingen

Department for Preclinical Imaging and Radiopharmacy

Werner Siemens Imaging Center

Roentgenweg 13

D-72076 Tuebingen

Germany

Telephone: +49 7071 29 87699

Fax: +49 7071 29 4451

e-mail: Jonathan.Disselhorst@med.uni-tuebingen.de

**First author:**

Prateek Katiyar, M.Sc.

University of Tuebingen

Department for Preclinical Imaging and Radiopharmacy

Werner Siemens Imaging Center

Roentgenweg 13

2
D-72076 Tuebingen

Germany

Telephone: +49 7071 29 83639

Fax: +49 7071 29 4451

e-mail: prateek.katiyar@med.uni-tuebingen.de

Financial support:

Grant support by: European Research Council, grant no. 323196

German Ministry for Education and Research/Bundesministerium für Bildung und Forschung (BMBF), grant number 0316186E

Eberhard Karls University Tuebingen (Evaluation of Tumor Heterogeneity Using Clustering of Multi-Modality Imaging Data, Fortuene 2131-0-0).
ABSTRACT

In this study we describe and validate an unsupervised segmentation algorithm for the assessment of tumor heterogeneity using dynamic $^{18}$F-FDG PET. The aim of our study was to objectively evaluate the proposed method and make comparisons with compartmental modeling parametric maps and standardized uptake value (SUV) segmentations using simulations of clinically relevant tumor tissue types. **Methods:** An irreversible two tissue compartmental model was implemented to simulate clinical and preclinical $^{18}$F-FDG PET time activity curves (TACs) using population based arterial input functions (80×clinical and 12×preclinical) and the kinetic parameter values of three tumor tissue types. The simulated TACs were corrupted with different levels of noise and used to calculate the tissue-type misclassification errors of spectral clustering (SC), parametric maps and SUV segmentation. The utility of the inverse noise variance (INV) and Laplacian score (LS) derived frame weighting schemes prior to SC was also investigated. Finally, the SC scheme with best results was tested on a dynamic $^{18}$F-FDG measurement of a mouse bearing subcutaneous colon cancer and validated using histology. **Results:** In the preclinical setup, the INV weighted SC exhibited the lowest misclassification errors (8.09% - 28.53%) at all noise levels in contrast to the LS weighted SC (16.12% - 31.23%), unweighted SC (25.73% - 40.03%), parametric maps (28.02% - 61.45%) and SUV (45.49% - 45.63%) segmentation. The classification efficacy of both weighted SC schemes in the clinical case was comparable to the unweighted SC. When applied to the dynamic $^{18}$F-FDG measurement of colon cancer, the proposed algorithm accurately identified densely vascularized regions from the rest of the tumor. In addition, the segmented regions and cluster-wise average TACs showed excellent correlation with the tumor histology. **Conclusion:** The promising results of SC mark its position as a robust tool for quantification of the tumor heterogeneity using dynamic PET studies. Since SC tumor segmentation is based on the intrinsic structure of the underlying data, it can be easily applied to other cancer types as well.

**Keywords:** Spectral clustering; tumor heterogeneity; compartmental modeling; $^{18}$F-FDG PET; SUV
INTRODUCTION

Tumors exhibit widespread genetic and phenotypic heterogeneity. The local tissue variability is known to mediate drug resistance and influence therapeutic efficacy (1). The magnitude of intra-tumor diversity is also linked with tumor aggressiveness and has shown to predict cancer mortality (2). The robust characterization of the tumor heterogeneity is not only an urgent requirement for precision medicine, but also for preclinical and pharmaceutical research (3).

The sensitivity and quantitative ability of positron emission tomography (PET) makes it a promising prognostic tool for cancer diagnosis and in vivo monitoring of therapy response. Accumulation of 18F-FDG in cancerous lesions is widely associated with tumor grade and prognosis (4,5). The most common clinical assessment of 18F-FDG is based on visual inspection and basic quantification of the standardized uptake value (SUV). Although the SUV as a metric is practical and easy to measure, it is vulnerable to numerous sources of variability (6). While static measures lack the ability to distinguish between non-phosphorylated and phosphorylated FDG, kinetic methods measure the complete aspects of the tracer distribution, providing vital information about glycolysis and blood flow. Kinetic modeling can play an especially essential role when evaluating the drug response of cancer patients with low pre-therapy FDG uptake, which results in poor sensitivity of the SUV and other static measures (7,8).

Despite the quantification benefits over static measures, kinetic methods such as compartmental modeling and graphical analysis have not been widely adopted, partly due to their reliance upon the acquisition of time activity curves (TACs) with low noise and a precise measurement of the arterial input function (AIF). Moreover, to improve signal to noise ratios, a common practice in dynamic PET studies is to perform region averaging (9) prior to compartmental modeling. Since compartmental modeling assumes the region of interest to be functionally homogeneous (10), user
defined delineations might lead to incorrect estimation of kinetic parameters in regions with tissue variability.

A voxel level analysis is essential to create a holistic profile of the spatial and temporal heterogeneity of cancerous lesions (11). Over the past decades, several segmentation methods have been proposed for the region wise analysis of PET images (12). Recently, one investigation has applied spectral clustering (SC) on dynamic PET data for brain image segmentation (13). Their investigations, however, lacked a histological validation. In this study we aimed to examine the suitability of SC in the segmentation of the tumor microenvironment. Through comprehensive simulations, we present an objective evaluation of SC and compare its robustness with the parametric maps and SUV segmentation. We also tested the proposed methodology in vivo on a mouse model of subcutaneous colon cancer with a histological validation.
MATERIALS AND METHODS

The widely accepted pharmacokinetic modeling tool COMKAT (14) was used to simulate \(^{18}\text{F-FDG}\) PET TACs. The complete details of the implemented compartmental model, preclinical experiments and histology are provided in the supplemental data.

Clinical and Preclinical Tissue Class Simulation

In order to simulate clinically relevant and comparable scenarios, the kinetic parameter values of different tissue classes were derived from Sugwara et al. (15). The authors studied 21 patients with primary germ cell tumors using \(^{18}\text{F-FDG}\) PET and reported the kinetic parameter values of three different tumor tissue classes, namely the viable tissue, the mature teratoma and necrosis. Since tumor tissue types were confirmed by histological findings, we extended the average kinetic parameter values of each tissue type as corresponding class representative. Likewise, the clinical AIF was selected from a population based AIF model (16). The study identified the parameters of the mathematical equations by fitting a three compartment blood pool model (17) on the arterial blood samples taken from 80 different patients. We contacted the authors to obtain the complete dataset because the published details were insufficient for simulations.

To extrapolate the clinical scenario into the preclinical setting, twelve 60 min dynamic \(^{18}\text{F-FDG}\) PET scans (4 mice × 3 scans) were acquired from 8-week old Naval Medical Research Institute nu/nu mice bearing subcutaneous Colo-205 tumors. The AIFs of all the measurements were approximated using a minimal blood sampling scheme (18). A two-tissue compartmental model was fitted to the mean time activity curve of each tumor, for each measurement. The obtained kinetic parameters from all 12 PET-scans provided realistic values of kinetic parameters observable in preclinical studies, which formed the basis to simulate the preclinical tumor tissue classes. Firstly, the averages of these kinetic parameters were used to simulate the viable tissue. Afterwards, the parameters of teratoma and
necrotic tissues were obtained by scaling the viable parameters to achieve the same parameter ratios (between different tissue classes) as in the clinical settings. The standard-deviations were chosen to match the mean/standard-deviation ratio of the respective clinical tissue type. All the animal experiments were performed in accordance to the German Animal Welfare Act, and local authorities approved all experimental protocols.

A total of 2000 TACs were sampled from a truncated Gaussian distribution (Table 1) for each tumor tissue class. The distributions were truncated to avoid sampling TACs with an unrealistic shape. The framing protocol was kept the same for both clinical and preclinical simulations: {30×2 s, 8×5 s, 8×10 s, 6×1 min, 5×2 min, 5×10 min}. For simplicity, throughout this paper, we refer to the simulated tumor tissue classes (viable, teratoma and necrosis) as class1, class2, and class3.

**Noisy TACs**

The noisy realizations of the simulated TACs were obtained by estimating the noise standard deviation for each time frame and distributing it log-normally to the noise free curve (9,19). The noise standard deviation for each frame \(i\) can be computed as follows:

\[
sd = \beta \sqrt{\frac{e^{-\lambda t_i \cdot ROI(t_i)}}{\Delta t_i}}.
\]

Where \(ROI(t_i)\) is the decay corrected activity concentration of the region of interest, \(e^{-\lambda t_i}\) is the decay un-correction factor, \(\lambda\) refers to the ratio \(\frac{\ln(2)}{\text{half-life}}\), \(\Delta t_i\) is the frame duration and \(\beta\) is a scale factor to limit the amount of noise within practical conditions. An illustrative example of noisy TACs can found in Supplemental Fig. 1.
SC

SC (20) uses Eigen structure of the affinity matrix and one of the classical clustering methods (e.g., k-means, fuzzy c-means, Gaussian Mixture Modeling) (21) to partition voxels into disjoint clusters. The affinity matrix \( W_{ij} \) of the dynamic PET data was computed as follows:

\[
W_{ij} = \begin{cases} 
    e^{-\|x_i - x_j\|^2/2\sigma^2} & \text{if } i \neq j \\
    0 & \text{otherwise.}
\end{cases}
\]

Here, \( \|x_i - x_j\| \) is the Euclidian distance between the TACs \( i \) and \( j \), and \( \sigma \) is the scale parameter of the Gaussian kernel. Subsequently, the affinity matrix \( W_{ij} \) was used to compute the normalized graph Laplacian using following expression:

\[
L = D_{inv} * W_{ij} * D_{inv}.
\]

Where, \( D_{inv} = D^{-0.5} \) and \( D \) is the diagonal matrix with \( d_i = \sum_{j=1}^{n} W_{ij} \) as the diagonal vector. To perform unsupervised clustering, the set of first \( k \) eigenvectors (corresponding to \( k \) largest eigenvalues (20)) of the normalized graph Laplacian was fitted using Gaussian Mixture Modeling. Throughout the study, we used the first 6 eigenvectors \( (k = 6) \) of the normalized Laplacian matrix and set \( \sigma \) equal to 40 and 55 for segmentation of preclinical and clinical TACs respectively. The scale was chosen experimentally, based on the misclassification error of the method on the noise free TACs. The same scale was used for segmentation of the preclinical example, but we could determine that segmentation was very robust to the choice of \( \sigma \).
PET Frame Weighting

In this study, the performance of two different weighting schemes for SC was investigated. In the first case, weights for each frame were set equal to the inverse of the noise variance (INV) of the respective frame, thus, dependent on frame length and total amount of activity in that specific frame. While in the second scheme, weights were derived from the Laplacian scoring (LS) algorithm (22). In the end, the weighted SC scheme with the best results (for preclinical simulations) was applied on the experimental data.

Clustering Comparisons

The clustering potential of SC was tested on the simulated data over varying levels of noise. The proposed methodology was also compared with SUV and parametric maps segmentation. In the former case, the average of the last two frames of the simulated dataset was clustered using $k$-means and in the latter case, the estimated kinetic parameters ($K_1$, $k_2$, $k_3$ and $K_4$) were segmented into three tissue classes using $k$-means and SC.

Evaluation Metrics

The percentage kinetic parameter estimation error ($\varepsilon$) was defined as:

$$
\varepsilon (\%) = \begin{cases} 
- \frac{K_{\text{estimated}}}{K_{\text{true}}} + 1 & \text{true} \leq \text{estimated} \\
\frac{K_{\text{estimated}}}{K_{\text{true}}} - 1 & \text{true} > \text{estimated.}
\end{cases}
$$

Where, $K_{\text{estimated}}$ is the estimated and $K_{\text{true}}$ is the true value of the compartmental modeling rate constant. The misclassification error was defined as follows:
\[
\frac{\sum_{i=1}^{N_{TAC}} \sum_{j=1}^{K} I(O_{i,j}, T_{i,j})}{N_{TAC} \times K}.
\]

Where \(O_{i,j}\) is the output and \(T_{i,j}\) is the true label of the TAC \(i\) from class \(j\), \(N_{TAC}\) represents the total number of TACs in each class and \(K\) is equal to the number of tumor tissue types i.e., 3. The indicator variable \(I\) is given as:

\[
I = \begin{cases} 
1 & \text{if } O_{i,j} \neq T_{i,j} \\
0 & \text{otherwise.}
\end{cases}
\]
RESULTS

Examples of simulated TACs of class1, class2, class3, and corresponding AIF for clinical and preclinical scenarios are shown in Fig. 1. To assess the influence of selected framing and the bias introduced by COMKAT, noise free curves were fitted using their respective AIF. The interquartile range and median ε for $K_1$, $k_2$, $k_3$ and $K_i$ for preclinical and clinical simulations are reported in Table 2.

Noise Evaluation

Fig. 2 shows the absolute ε for noisy preclinical TACs with different levels of log-normally distributed noise ($\beta = 0.1-1.5$). Amongst all, $k_2$ and $k_3$ showed the highest deviations from the true parameter values. Moreover, the errors in $k_2$ and $k_3$ also propagated to $K_i$. A similar tendency was seen in the case of noisy clinical TACs (Supplemental Fig. 2). Although the ε for $k_3$ and $K_i$ in the clinical case carried less variability than those in the preclinical settings.

The segmentation ability of different clustering methods for noisy preclinical TACs is shown in Fig. 3. While the INV weighted SC exhibited the lowest misclassification error, both the weighted and unweighted SC techniques outperformed other clustering schemes. Fig. 3 also depicts the misclassification errors obtained after clustering the SUV and estimated kinetic parameters. Up to moderate noise levels ($\beta < 0.7$), $k$-means and SC applied on the estimated kinetic parameters yielded lower errors in comparison to clustering the SUV, signifying the efficacy of dynamic measures over the static ones. Supplemental Fig. 3 shows the aforementioned clustering results for the clinical scenario. At low noise levels ($\beta < 0.5$), SC on the estimated kinetic parameters displayed the highest accuracy, but became worse with a gradual increase in noise. Overall for clinical simulations, the misclassification error of LS weighted SC remained most steady at all noise levels.

Supplemental Fig. 4 shows the ground truth and clustering affinity matrices for the noise free preclinical TACs (shown in Fig. 1B). It is clearly visible that the clustering solution retains the
approximate block diagonal structure of the original affinity matrix. Here, the clustering solution corresponds to the INV weighted SC of the simulated noise free preclinical TACs. The grid lines in Supplemental Fig. 4B give an impression as to the extent of over estimation of class3 and respective underestimation in class1.

Example

Fig. 4 shows the segmentation result of INV weighted SC on an 18F-FDG measurement. The algorithm effectively identified the densely vascularized regions (depicted with blue and red color) from the rest of the tumor (green cluster). The segmented regions were visually validated by CD-31 histology of the tumor section (Fig. 4A). The affinity matrix of the aforementioned clustering solution is shown in Supplemental Fig. 5B. The average TACs of well perfused areas also showed a significantly higher uptake than that of the rest of the tumor (Supplemental Fig. 5C). The parametric maps of this tumor are presented in Fig. 5. It also shows an 18F-FDG PET image exhibiting the tumor uptake in the last 20 min of the scan. The outcome of segmenting the tumor parametric maps using SC is shown in Supplemental Fig. 5D-F. It is evident from Supplemental Fig. 5B that clustering the tumor TACs yielded compartments with high intra-cluster similarity, while the uncertainties in the parametric maps resulted in poor segmentation of the tumor with low within-cluster similarity (Supplemental Fig. 5D-F).
DISCUSSION

This study shows the potential of spectral clustering for the assessment of tumor heterogeneity using dynamic \(^{18}\)F-FDG PET data. It also contrasts SC with the widely utilized two tissue compartmental model and the SUV, using dynamic PET simulations of clinically relevant tumor tissue types. The clinical tissue classes were duplicated in preclinical setting and studied for different levels of noise. A meaningful comparison of the proposed algorithm with compartmental modeling was carried out by fitting the noisy TACs and subsequently clustering the estimated kinetic parameters using \(k\)-means and SC. Furthermore, as a proof of principle we also applied the suggested method to an \textit{in vivo} mouse model of colon cancer and validated it with histology. Recently, the value of unsupervised segmentation has been shown in a translational study (23). The promising results of SC on the simulated datasets as well as on an \textit{in vivo} mouse model strongly indicate its potential for dynamic \(^{18}\)F-FDG PET clinical investigations.

A precise characterization of the tumor microenvironment requires a robust voxel level analysis. However, the variability of kinetic rate constants with the amount of noise and distortions in AIF (9,24) indicates the shortcomings of compartmental modeling for a voxel wise analysis. Although clustering the estimated kinetic parameters in the preclinical case produced more accurate results than clustering the SUV (\(\beta < 0.7\)), the misclassification error of the INV weighted SC was lower than that of any of the other schemes. In clinical simulations, SC applied on the estimated kinetic parameters seemed promising at low noise levels, but failed to distinguish tumor tissue types accurately as the TACs became noisier. It should be noted that \(k\)-means and SC errors on the estimated kinetic parameters reflect the best case scenario for compartmental modeling based tumor tissue segmentation, since the noisy TACs were modeled using their respective true AIFs (without any shape distortions). Uncertainties in AIF are most likely to introduce adverse effects on kinetic parameter estimation and consequently in parametric maps based tumor tissue segmentation. The poor predictive ability of the SUV in both preclinical and
clinical settings was due to the considerable overlap in the last time points of the TACs of all three tumor tissue types. This shows that the faster static PET acquisition comes at the cost of vital physiological information, which can play a principal role in probing intra-tumoral heterogeneity. The errors caused by noise in kinetic modeling on the other hand, can be minimized to a moderate extent by first utilizing the proposed algorithm for region segmentation and later estimating the kinetic parameters from the averaged TACs.

In dynamic PET imaging, early, middle, and late frames capture different kinetics of the TAC. However, due to non-uniform frame durations and different activity concentration levels they are also affected by varying levels of noise. Thus, while clustering the simulated TACs, we compared the efficacy of two different frame weighting schemes: INV and LS. While the former scheme intuitively favors frames with higher signal to noise ratio, the latter one exploits the intrinsic structure of the high dimensional data. In the analysis of preclinical simulations the INV weighted SC performed marginally better than the LS weighted SC; the opposite was true in the case of clinical simulations.

Some of the results presented in this paper may slightly vary with a different choice of frame sampling schedule. For example, longer early frames might increase the robustness of kinetic parameter estimates at the expense of faster early kinetics. Likewise, the rebinning will also influence the misclassification errors of different clustering schemes. Since this can be an independent study on its own, we did not optimize the simulations for the best framing schedule. Similar considerations apply to different tracer infusion protocols. Furthermore, in order to be consistent with Sugwara et al. (15), the two tissue compartmental model was implemented with $k_4$ and $F_b$ (fractional blood volume) equal to 0. Although the tissue types identified in colon cancer were different from the simulated tissue classes (except for the viable), the synthetic TACs enabled a thorough objective evaluation of the proposed technique. Moreover, since SC tumor segmentation is based on the intrinsic structure of the underlying data, it can be easily applied to other cancer types as well.
The number of clusters in the example in Fig. 4 was determined based on the visual inspection of the data and solution affinity matrices for different number of clusters. Significant off-diagonal similarity between the red and blue clusters was evident from the similar average TACs of the respective regions (Supplemental Fig. 5C). While the blue cluster corresponds well to regions with high vessel density, the red cluster appears to the periphery of the blue regions, resulting in similar uptake patterns. It is important to mention that the histology was rigidly registered with the imaging and we did not perform any non-rigid registration between the two. Although the tumor was carefully partitioned into two parts parallel to the transversal field of view, imaging to histology registration remains non-trivial due to the substantial differences in resolution (mm vs. µm). Additionally, during the dehydration process the tissue sections undergo a series of non-deterministic affine deformations, which cannot be corrected using rigid transformations. However, by sectioning the tumor along the reference (imaging) plane and keeping a track of its orientation, the errors in the manual registration can be minimized (25).

Unlike K-means, SC does not make any assumptions about the shape of the clusters. The efficacy of SC mainly lies in the change of representation (from abstract data points to points in the feature space), which enhances the segregation tendency of the input data. It should be noted that the optimal SC solution depends upon the number of chosen eigenvectors from the normalized graph Laplacian. In ideal scenarios, the top \( k \) eigenvectors corresponding to the \( k \) largest eigenvalues of the normalized Laplacian matrix (where \( k = \) number of biological classes) contain the class discriminative information (20). However, because of the complex microenvironment, resolution limit and large statistical noise, compartments in oncological dynamic PET studies often display similar tracer uptake patterns. To a certain extent, these perturbing effects can be dealt with by choosing a larger number of eigenvectors than the potential number of clusters. Throughout our study, we utilized 6 eigenvectors to segment the dynamic PET data (simulated and measured) into relevant biological compartments. It has been shown that a prior eigenvector selection can further enhance the clustering stability (26), but we did not
explore any such possibility. Additionally, the choice of graph Laplacian can also affect the outcome of SC. As suggested in literature (20), we used the normalized graph Laplacian rather than the un-normalized one. Also, we did not notice any difference in the performance of two previously established normalized graph Laplacians.

A clear limitation of this study is the lack of clinical experimental data, however accurate alignment of histology to imaging in a clinical setting is difficult to achieve, making validation of intratumoral tissue classes challenging. In preclinical studies, this alignment can be more easily performed. Yet, Fig. 4 only presents a qualitative comparison of the segmented tumor compartments with the histology. Future preclinical studies will include an automated non-rigid imaging to histology co-registration to provide reliable quantification of intratumoral heterogeneity. Since PET scanners have a finite spatial resolution, tissue inhomogeneities occurring at the cellular level cannot be observed and analyses are limited to large-scale heterogeneity. Information about variations at this scale has clear potential, for example, in radiotherapy for dose painting and as a basis in image-guided biopsy procedures.

To the best of our knowledge, this is the first study investigating the feasibility of SC for the assessment of the tumor microenvironment incorporating exhaustive dynamic PET simulations and augmented by real data with histological validation. SC exploits the temporal characteristics of dynamic studies and uses high dimensional embedding (27) to effectively segment the tumor into distinct biological compartments. This could play an instrumental role in in vivo cancer studies, as the tumor microenvironment stems from complex genetic alterations and phenotypic interactions, which might not be readily discernible using the existing methods for analyzing dynamic PET measurements.
CONCLUSION

We have shown the feasibility of SC for the segmentation of 4-D dynamic PET tumor images. The proposed technique showed superior performance in comparison to the SUV and parametric maps based segmentation of tumor tissue variability. Overall, SC can be used as a potential tool for the voxel level characterization of the tumor microenvironment.

ACKNOWLEDGEMENTS

We gratefully thank Dennis Vriens for providing the clinical AIF data and many useful comments on the manuscript.

DISCLOSURE

This study was financially supported by European Research Council (ERC), grant agreement number 323196, German Ministry for Education and Research/Bundesministerium für Bildung und Forschung (BMBF), grant number 0316186E, and Eberhard Karls University Tuebingen (Evaluation of Tumor Heterogeneity Using Clustering of Multi-Modality Imaging Data, Fortuene 2131-0-0).
REFERENCES


FIGURE 1. Class-wise simulated TACs and the corresponding AIF for clinical (A and C), and preclinical (B and D) scenarios. Kinetic parameters for each class were sampled from truncated Gaussian distributions. Shaded regions depict the distribution of time activity curves up to unit standard deviation of the respective tumor tissue type.
FIGURE 2. Absolute $\varepsilon$ for preclinical simulations with an increase in the amount of noise ($\beta$) for $K_1$ (A), $k_2$ (B), $k_3$ (C), and $K_i$ (D). The boxes depict the interquartile range and whiskers represent the 10th and 90th percentiles of the data.
FIGURE 3. Misclassification error of various clustering schemes for preclinical simulations with an increase in the amount of noise ($\beta$).
FIGURE 4. (A) CD31 stained histology of a representative tumor; the four insets (scale in µm) illustrate high vessel density areas. (B) Segmentation of the tumor into three clusters by applying SC on the dynamic $^{18}$F-FDG PET data. The matched clusters are marked as a, b, c, and d in Fig. A and B respectively.
FIGURE 5. (A) Left to right: $K_1$, $k_2$ and $k_3$ maps of the tumor shown in Fig. 4. (B) Left to right: $K_i$ map calculated using the parametric maps in A and $^{18}$F-FDG uptake in the tumor in the last 20 min of the scan.
TABLE 1. Summary of the kinetic parameters (mean ± standard deviation) and corresponding truncation limits used for the simulation of preclinical and clinical tumor tissue classes.

<table>
<thead>
<tr>
<th>Kinetic parameter</th>
<th>Class1 Preclinical</th>
<th>Class2 Preclinical</th>
<th>Class3 Preclinical</th>
<th>Class1 Clinical</th>
<th>Class2 Clinical</th>
<th>Class3 Clinical</th>
<th>Truncation limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>K₁</td>
<td>0.138 ± 0.043</td>
<td>0.123 ± 0.033</td>
<td>0.114 ± 0.026</td>
<td>0.045 ± 0.006</td>
<td>0.036 ± 0.005</td>
<td>0.01-1.0</td>
<td></td>
</tr>
<tr>
<td>k₂</td>
<td>0.116 ± 0.136</td>
<td>0.180 ± 0.069</td>
<td>0.105 ± 0.025</td>
<td>0.176 ± 0.042</td>
<td>0.01-1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k₃</td>
<td>0.085 ± 0.056</td>
<td>0.014 ± 0.008</td>
<td>0.005 ± 0.002</td>
<td>0.001-1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 2. Kinetic parameter estimation errors obtained after fitting the preclinical and clinical noise free TACs using respective AIFs.

<table>
<thead>
<tr>
<th>Kinetic parameter</th>
<th>Kinetic parameter estimation error (ε) %</th>
<th>Preclinical</th>
<th>Clinical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>IQR</td>
<td>Median</td>
</tr>
<tr>
<td>K₁</td>
<td>-0.003</td>
<td>(-0.020, 0.012)</td>
<td>0.012</td>
</tr>
<tr>
<td>k₂</td>
<td>-0.040</td>
<td>(-0.180, 0.071)</td>
<td>0.050</td>
</tr>
<tr>
<td>k₃</td>
<td>-0.121</td>
<td>(-0.649, 0.213)</td>
<td>0.049</td>
</tr>
<tr>
<td>Kᵢ</td>
<td>-0.033</td>
<td>(-0.385, 0.091)</td>
<td>0.000</td>
</tr>
</tbody>
</table>
SUPPLEMENTARY MATERIALS AND METHODS

Two-tissue Compartmental Modeling

A simplified two tissue compartmental model for $^{18}$F-FDG is shown in Supplemental Fig. 6, where $C_p(t)$, $C_1(t)$ and $C_2(t)$ are the time varying plasma, free and bound tracer activity concentrations (Bq/mL); $K_1$, $k_2$, $k_3$ and $k_4$ are the kinetic rate constants which control the rate of tracer exchange between compartments. In order to simulate TACs, an irreversible two tissue compartmental model (i.e., $k_4 = 0$) was implemented. Additionally, the vascular fraction ($F_v$) was assumed to be zero. The equations for this model can be written in the following manner:

$$\frac{dC_1(t)}{dt} = K_1 C_p(t) - k_2 C_1(t) - k_3 C_1(t)$$

$$\frac{dC_2(t)}{dt} = k_3 C_1(t)$$

$$C_{model}(t) = C_1(t) + C_2(t).$$

Here, $C_{model}(t)$ is the observed activity concentration of the target tissue. The model parameters ($K_1$, $k_2$ and $k_3$) are estimated by solving the ordinary differential equations and minimizing a weighted least square objective function. The net influx rate ($K_i$) was computed as:

$$K_i = \frac{K_1 \cdot k_3}{k_2 + k_3}.$$
Preclinical Experiments

Six-week-old Naval Medical Research Institute nu/nu mice (n = 4) were ordered from Charles River, Germany and allowed to acclimatize in the on-site animal vivarium before being subcutaneously injected with $4.5 \times 10^6$ Colo-205 tumor cells on the right hind leg. The tumor size as well as normal social activity of animals was monitored during the entire study. When the tumors were palpable and showed signs of vascularization with a minimum length of 5 mm, the imaging experiments were started. Before and during tumor inoculation and imaging experiments, mice spontaneously respired 1.5% isoflurane dissolved in 100% O$_2$ at a flow rate of 0.8 L/min in order to maintain a deep anesthesia. All experiments were carried out in a specific-pathogen-free environment.

Three line sources were placed on the animal holder in order to co-register PET and magnetic resonance imaging (MRI) data. The mice were placed with the tumor in the middle of the field-of-view of the Inveon dedicated small animal PET scanner (Siemens, Knoxville, TN, USA) and 12.0 MBq of $^{18}$F-FDG in 50 μL of 0.9% NaCl was injected in the tail vein at a flow rate of 0.5 mL/min using an automated syringe pump system (Harvard Apparatus, Holliston, MA, USA) directly after starting the PET acquisition. Scans were acquired using manufacturer-supplied software (Inveon Acquisition Workplace, version number 1.5.0.28) for 60 min and reconstructed using OSEM3D/FastMAP (reconstruction software version: 2.5, histogram version: 2.39 and re-binning version: 2.5) with the following framing: {10×2 s, 4×5 s, 2×10 s, 3×1 min, 3×2 min, 5×10 min}. Other reconstruction parameters were as follows: image zoom = 1, image matrix size = 256×256×159 with (0.39×0.39×0.80) mm$^3$ voxel sizes, OSEM3D Iterations = 2, MAP Iterations = 18, Beta = 0.05, Uniform set to Resolution and FastMAP setting on.

After the PET scans, the animal holder was moved to a 7 T Clinscan (Bruker BioScan, Ettlingen, Germany) small animal MRI while maintaining the position of the mouse. The following settings were used for the T2 weighted turbo spin echo sequence (T2se): repetition time (TR) = 3000 ms, echo time (TE) = 205 ms, echo train length = 161, image size = 256×160 and voxel size (mm$^3$) = 0.22×0.22×0.22.
PET images were co-registered to MR images using a marker-based, semi-automatic co-registration tool in PMOD 3.2 (PMOD Technologies, Zurich, Switzerland) and the T2tse images were used as an anatomical reference for drawing volumes of interest (VOIs) on each tumor. Special care was taken to exclude the skin of the mice during the VOI placement on the tumors in T2tse images. The voxel values along with the coordinates from each VOI for all PET and MRI measurements were exported and further processed in MATLAB (Mathworks, Natick, MA, USA).

**Histology**

The histology and immunohistochemistry of one tumor was obtained to validate the results of SC on $^{18}$F-FDG measurements. Following the dynamic PET scan, the mouse was sacrificed and a line was drawn on the tumor parallel to the transversal imaging plane. The tumor was removed using a scalpel and sectioned into two halves along the aforementioned line. The individual halves were kept in neutral buffered formaldehyde (4.7% by volume) and embedded into paraffin, before processing for staining. For histology, 3-5 µm-thick sections were cut and stained with haematoxylin and eosin. Immunohistochemical stainings with an anti-CD31 antibody (Abcam plc) were performed on an automated immunostainer (Ventana Medical Systems, Inc.) according to the company’s protocols with slight modifications. Appropriate positive and negative controls were used to confirm the adequacy of the staining. Only a rigid co-registration was performed between the histology and imaging data. The matching imaging slice was selected based on the visual alignment of the contours of the PET and histology image.
Supplementary figure 1. Preclinical noise free and noisy (for $\beta = 0.5$) TAC samples of class1, class2, and class 3.
Supplementary figure 2. Absolute $\epsilon$ for clinical simulations with an increase in the amount of noise ($\beta$) for $K_1$ (A), $k_2$ (B), $k_3$ (C), and $K_i$ (D). The boxes depict the interquartile range and whiskers represent the 10th and 90th percentiles of the data.
Supplementary figure 3. Misclassification error of various clustering schemes for clinical simulations with an increase in the amount of noise (β).
Supplementary figure 4. Ground truth (A) and clustering (B) affinity matrices of the preclinical noise free TACs (shown in Figure 1B). The three block diagonal matrices (top to bottom) depict the intra-class similarity of class 1, 2, and 3. The rest of the two off-diagonal block matrices in each row display the inter-class similarity between the annotated classes. As the TACs of class 2 and class 3 were relatively similar in shape (Fig. 1B), the inter-class similarity between these two clusters was also higher. On the other hand, despite significant overlap, the inter-class similarity between class 1 and class 2 was relatively lower, primarily due to the differences in the shape of the simulated TACs.
Supplementary figure 5. (A) SC segmented image of a representative tumor with three clusters obtained using dynamic $^{18}$F-FDG PET data. (B) Affinity matrix of the entire tumor volume computed using the clustering solution in A. The comparable cluster population of the green and blue regions is due to the fact that the end slices of the tumor are densely vascularized in contrast to the center ones. (C) Averaged time activity curves of the respective clusters. While all three clusters depict high intra-class similarity, the red and blue clusters also contain high inter-class similarity. As the red cluster appears on the periphery of the blue cluster, both the regions also have similar average TACs. (D) Segmentation of the same tumor using the parametric maps (shown in Fig. 5) and SC. (E) Affinity matrix of the entire tumor volume computed using the clustering solution in D and PET TACs. (F) Averaged time activity curves of the respective clusters.
Supplementary figure 6. A two tissue compartmental model for $^{18}$F-FDG PET. The dashed box represents the observed activity concentration of the region.
Spectral Clustering predicts tumor tissue heterogeneity using dynamic $^{18}$F-FDG PET: a complement to the standard compartmental modeling approach

Prateek Katiyar, Mathew R Divine, Ursula Kohlhofer, Leticia Quintanilla-Martinez, Bernhard Schölkopf, Bernd J Pichler and Jonathan A Disselhorst

*J Nucl Med.*

Published online: November 3, 2016.

DOI: 10.2967/jnumed.116.181370

This article and updated information are available at:

http://jnm.snmjournals.org/content/early/2016/11/02/jnumed.116.181370

Information about reproducing figures, tables, or other portions of this article can be found online at:

http://jnm.snmjournals.org/site/misc/permission.xhtml

Information about subscriptions to JNM can be found at:

http://jnm.snmjournals.org/site/subscriptions/online.xhtml

*JNM* ahead of print articles have been peer reviewed and accepted for publication in *JNM*. They have not been copyedited, nor have they appeared in a print or online issue of the journal. Once the accepted manuscripts appear in the *JNM* ahead of print area, they will be prepared for print and online publication, which includes copyediting, typesetting, proofreading, and author review. This process may lead to differences between the accepted version of the manuscript and the final, published version.

*The Journal of Nuclear Medicine* is published monthly.

SNMMI | Society of Nuclear Medicine and Molecular Imaging

1850 Samuel Morse Drive, Reston, VA 20190.

(Print ISSN: 0161-5505, Online ISSN: 2159-662X)

© Copyright 2016 SNMMI; all rights reserved.