

Aus der Klinik für Strahlentherapie und Radioonkologie
der Medizinischen Fakultät Mannheim
(Direktor: Prof. Dr. med. Frederik Wenz)

**[¹⁸F]FET-PET brain image segmentation using *k*-means:
Evaluation of five cluster validity indices**

Inauguraldissertation
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Bedor Abualhaj
aus
Amman, Jordanien

2017

Dekan: Herr Prof. Dr. med. Sergij Goerd

Referent: Herr Prof. Dr. med. Frederik Wenz

Table of Contents

List of Abbreviations.....	v
List of Figures.....	vi
List of Tables.....	viii
1. Introduction.....	1
1.1 Positron emission tomography	2
1.1.1 General basic principles of PET imaging (physics)	2
1.1.2 PET radiotracers	3
1.1.3 Interaction of gamma rays with matter.....	4
1.1.4 Photon detection and data acquisition	5
1.1.5 Classification of detected events in PET	5
1.1.6 Image reconstruction	6
1.1.7 Static and dynamic PET images	7
1.2 PET image segmentation using cluster analysis	7
1.2.1 <i>k</i> -means clustering	8
1.2.2 Cluster validity indices.....	10
1.2.3 Validity indices used for PET image segmentation	10
1.3 Aims	12
1.4 Structure of the thesis	12
2. Materials and Methods	13
2.1 Patient and data acquisition	13
2.2 PET data preprocessing	14
2.3 Clustering using the <i>k</i> -means algorithm	15
2.4 Reproducibility of <i>k</i> -means objective function value	15
2.5 Cluster validity indices.....	16
2.6 Determination of the needed number of replications for each validity index	19
2.7 Evaluation of validity indices for the general requirements	20

3. Results22
3.1 Reproducibility of *k*-means objective function value.....24
3.2 Cluster validity index requirements24
4. Discussion41
5. Conclusions48
6. References49
7. Curriculum Vitae53
8. Publications and Abstracts55
9. Acknowledgements.....56

List of Abbreviations

ACD	Annihilation Coincidence Detection
CT	Computed Tomography
CV	Coefficient of Variation
^{18}F	Fluorine-18
FDG	Fluoro-deoxy-glucose
FET	Fluoro-ethyl-L-tyrosin
LOR	Line of Response
MRI	Magnetic Resonance Imaging
OFV	Objective Function Value
PET	Positron Emission Tomography
PMT	Photomultiplier Tubes
ROI	Region of Interest

List of Figures

FIG.1. Reconstructed dynamic [¹⁸ F]FET-PET images for a patient had oligodendroglioma. Note that concentrations higher than 6000 (Bq/cc) were set to 6000 (Bq/cc).....	22
FIG.2. Reconstructed dynamic [¹⁸ F]FET-PET images for a patient had glioblastoma. Note that concentrations higher than 21415 (Bq/cc) were set to 21415 (Bq/cc).....	23
FIG.3. Coefficient of variation (<i>CV</i>) of the <i>k</i> -means objective function value (<i>OFV</i>) using 100 random centroid initialization replications (<i>RCI</i> ₁₀₀) to evaluate the reproducibility of <i>k</i> -means, for number of clusters from 2 to 50.....	23
FIG.4. Optimal number of clusters suggested by the AIC as a function of the number of voxels (range 300 – 30000) showing the violation of requirement 2 (TABLE II; note the logarithmic scale for the x-axis).....	26
FIG.5. Cluster validity indices plot for the number of clusters from 2 to 50 with 100 different random initial centroids of the patient shown in FIG.1. At the optimal number of clusters, defined by WB, <i>I</i> (note the logarithmic scale for the y-axis), modified Dunn’s and Silhouette indices, the coefficients of variation were 7×10^{-5} , 2×10^{-1} , 2×10^{-4} and 2×10^{-4} , respectively.....	28
FIG.6. Cluster validity indices plot for the number of clusters from 2 to 50 with 100 different random initial centroids of the patient shown in FIG.2. At the optimal number of clusters, defined by WB, <i>I</i> (note the logarithmic scale for the y-axis), modified Dunn’s and Silhouette indices, the coefficients of variation were 3×10^{-2} , 8×10^{-16} , 2×10^{-2} and 3×10^{-3} , respectively.....	29
FIG.7. The optimal number of clusters is defined for a given range from 2 to 50 according to (a) WB, (b) <i>I</i> (note the logarithmic scale for the y-axis), (c) modified Dunn’s and (d) Silhouette validity indices for the patient shown in FIG.1. The replication with the minimal objective function value was selected from 100 replications.....	30
FIG.8. The optimal number of clusters is defined for a given range from 2 to 50 according to (a) WB, (b) <i>I</i> (note the logarithmic scale for the y-axis), (c) modified Dunn’s and (d) Silhouette validity indices for the patient shown in FIG.2. The replication with the minimal objective function value was selected from 100 replications.....	31

FIG.9. *k*-means segmentation (upper row) and superimposed tumor contours (lower row) of the [¹⁸F]FET-PET kinetics of the patient with an oligodendroglioma (FIG.1) for the optimal number of clusters *NC*=5 given by WB and *I* indices (left column) and the optimal number of clusters *NC*=3 given by the modified Dunn's and Silhouette indices (right column). Different clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0. The contour was detected using the 'log' edge algorithm (Laplacian operator) of MATLAB (R2015b, MathWorks, Natick, Massachusetts, USA) software. The obtained contours were checked by an experienced physician for plausibility. Non-tumor cluster areas such as veins, arteries and the sinus cavernous or enrichment of the mucous membranes of the nasal sinuses in the [¹⁸F]FET-PET brain images are not contoured.....33

FIG.10. *k*-means segmentation (upper row) and superimposed tumor contours (lower row) of the [¹⁸F]FET-PET kinetics of the patient with glioblastoma (FIG.2) for the optimal number of clusters *NC*=6 given by WB-index (left column) and the optimal number of clusters *NC*=4 given by *I* index (middle column) and the optimal number of clusters *NC*=2 given by the modified Dunn's and Silhouette indices (right column). Different clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0.....34

FIG.11. TACs corresponding to the tumor clusters (seen in FIG.9) based on the optimal number of clusters suggested by the WB-index or *I*-index (a) and (b) and by the optimal number of clusters suggested by modified Dunn's index or Silhouette index (c). The numbers of voxels were 597, 110, and 601, for tumor clusters 1, 2 and 3, respectively.....35

FIG.12. TACs corresponding to the tumor clusters (seen in FIG.10) based on the optimal number of clusters suggested by the WB-index (a) and (b), the optimal number of clusters suggested by *I*-index (c) and (d) and the optimal number of clusters suggested by modified Dunn's index or Silhouette index (d) and (f). The numbers of voxels were 175, 115, 498, 128 and 414 for tumor clusters 1, 2, 3, 4 and 5, respectively.....36

FIG.13. The top row shows an example for three neighboring slices having tumor in the last frame, (a) slice 25 (b) slice 26 (c) slice 27 taken from patient shown in (FIG.1). *k*-means segmentation result can be seen in the bottom row for the three neighboring slices with the optimal number of clusters, (a) 4 (b) 5 (c) 5, given by WB and *I* validity indices. Clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0. The slices **b** and **c** have one cluster more than slice **a** due to heterogeneity of tumor in the selected plane.....39

FIG.14. The top row shows an example for three neighboring slices having tumor in the last frame, (a) slice 42 (b) slice 43 (c) slice 44 taken from patient shown in (FIG.2). *k*-means segmentation result can be seen in the bottom row for the three neighboring slices with the optimal number of clusters, (a) 5 (b) 6 (c) 5, given by WB index. Clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0. The slices **b** has one additional cluster more than slices **a** and **c**.....40

List of Tables

TABLE I. Cluster Validity indices.....17

TABLE II. General requirements for the optimal number of clusters determined based on an adequate cluster validity index.....25

TABLE III. Comparison of the optimal number of the clusters for three neighboring slices as defined by different cluster validity indices. Note: From first row to the eighth row, the first three numbers in each slot are the optimal number of clusters corresponding to three neighboring slices separately. The number in the bracket is the optimal number of clusters corresponding to the same three slices together.....38

1. Introduction

The accurate, reproducible and precise identification of the spatial distribution for tumor regions and adjacent healthy tissues is a required step before radiotherapy treatment planning.¹ After identifying these regions, physicians prescribe absorbed doses to the tumor tissues and define maximum doses to the organs at risk. Traditionally, the delineation of the target regions and the surrounding organs at risk is defined based on anatomical images obtained from computed tomography (CT).²⁻⁶

However, not all spatial regions inside the tumor can be identified or might be visible on a CT scan. In addition, the effect of the prescribed absorbed dose to the tumor regions depends on specific properties of the tumor such as its oxygenation status and the proliferation rate of the tumor cells. Therefore, it is useful to include further imaging modality techniques such as positron emission tomography (PET) to acquire deeper physiological insights about the heterogeneity of the tumor and thus to support treatment planning.⁷⁻¹³ In fact, tumor heterogeneity is a required information for treatment planning as different metabolic structures within the tumor may have different responses to radiotherapy and this requires adjusting the prescribed absorbed dose for each individual region.¹⁴

In this chapter contents are organized in two sections. The physical principle of positron emission tomography (PET) is provided and described in the first section. Also, the section introduces a review of the formation process of PET images which starts from patient injection, photon detection and data collection and ends with image reconstruction. In the second section, the focus is on PET image segmentation using *k*-means, followed by investigating and checking the

reproducibility of this algorithm and evaluating the accuracy of PET image segmentation using five cluster validity indices.

1.1 Positron Emission Tomography

PET is a functional nuclear medicine imaging technique. It is considered as an important tool in radio-oncologic treatment planning, as it allows studying the relevant molecular processes in the tumor cells non-invasively.¹⁵⁻¹⁸ It has several applications in research, before and after radiotherapy treatment planning. Among others, it can be used for tumor diagnostics, tumor staging and restaging, optimizing the radiotherapy by adjusting the prescribed dose according to specific properties of different sub-regions within the tumor as well as by monitoring and evaluating the tumor response to radiotherapy treatment planning.^{19, 20} Additionally, it is used to differentiate recurrence from radiation necrosis.²⁰

1.1.1 General basic principles of PET imaging (physics)

When a patient is required to take a PET in clinical routine, the patient is injected with small amount of labeled molecules or compounds with positron emitting radioisotopes (radiotracer). After a specific time and appropriate uptake period, the radiolabeled tracer subsequently is transported and distributed inside the human body and accumulates particularly in the tumor regions.^{10, 21, 22} PET images are acquired based on the emission of positron particle (the anti-particle of electron) from the injected radiolabeled tracer during the physical decay of the attached radioisotope. The emitted positron travels a short distance (the positron range) which not only affected by its initial energy but also by the properties and characteristics of the surrounding

tissues including electron density and atomic number. At the end of its path, it combines with an electron by annihilation in which their rest masses are converted into gamma photons. Consequently, two 511 keV gamma photons travelling in opposite direction are emitted within the human body. These high energetic gamma photons might be located and registered in coincidence (simultaneously) by PET detectors which are arranged in a ring around the patient.^{10, 21, 22}

To support the treatment planning technique, the PET is commonly integrated with a CT or magnetic resonance tomography (MRT). The combined modalities provide both anatomic and functional information about different metabolic processes in the human body.⁸⁻¹³ Various positron emitting isotopes are available which can be incorporated into biological molecules without changing their chemical properties. Moreover, their half-life is relatively suitable to avoid high dose to the patients and to minimize the radiation risk while the required diagnostic information still can be obtained.

1.1.2 PET radiotracers

A radiotracer is a biological compound or molecule which is labelled with one of various available radioisotopes without changing their chemical properties. Through the radioactive decay including emitting particles, the nucleus changes its compositions and reaches a more stable state.^{10, 21, 22} The most frequently used radioisotopes in PET are ^{18}F , ^{12}C , ^{13}N , which have too many protons compared to the number of neutrons. Thus, to reach the stability, the decay mode of these isotopes is positron emission. Among all, ^{18}F is the best currently used radioisotope due to its suitable half-life (109.8 min) and its suitable mean range (0.6 mm). It is

used to produce for example [^{18}F]fluoro-deoxy-glucose ([^{18}F]FDG) and [^{18}F]fluoro-ethyl-L-tyrosine [^{18}F]FET.

Depending on specific clinical application and the molecular processes being investigated, the patient is injected with one type of several available radiolabeled tracers used for PET. Most frequently, [^{18}F]FDG is used in oncology, because it measures glucose metabolism in different tissues, which is enhanced in tumors.^{23, 24} However, the high accumulation of this tracer in brain or due to inflammation is problematic.²⁵ A more specific radiotracer for identifying tumor regions is [^{18}F]FET, which is an amino-acid radiolabeled tracer. This tracer accumulates in brain tumors rather than normal brain tissues due to the higher number of amino acid transporters in tumor cells.²⁶⁻²⁸

1.1.3 Interaction of gamma rays with matter

After the radioactive decay and the annihilation radiation process, photons are emitted within the human body which can be scattered and absorbed when passing through the tissues. The possible interaction mechanisms with the matter are affected by the energy of the incoming photons. The photoelectric effect and Compton scattering interaction predominate when the photon energy is in the (K-eV) range. For higher photon energies (range 1.02 M-eV and more), pair production becomes important. Furthermore, the probability of these interactions depends on the atomic number and the electron density of the matter in such a way that the probability of Compton scattering increases along with decreasing the atomic number of the material, whereas, for dense materials with higher atomic numbers, the photoelectric effect is much more significant.^{10, 21, 22, 29,}

1.1.4 Photon detection and data acquisition

The PET detector consists of a dense crystalline scintillator material. When the produced high energetic two photons reach this material, visible light is emitted. Then, the produced visible light is converted into electrical current by the photomultiplier tube (PMT) which is connected with PET detectors. The main components of this tube are photocathode, anode, and several dynodes. When the incident visible light strikes the photocathode, electrons are injected which are accelerated from the photocathode to the first dynode by the high potential difference. When hitting the first dynode, more electrons are emitted through secondary emission process. These emitted electrons are in turn accelerated to the second dynode. This process is repeated until reaching the last dynode. After being amplified through (PMT), the multiplied secondary electrons are attracted to the anode, which in turn transfers and delivers the electron current to an external electronic circuit involving preamplifier, amplifier, pulse height analyzer and other subsequent electronics.^{10, 21, 22, 29, 30}

1.1.5 Classification of detected events in PET

PET scan is based on the annihilation coincidence detection (ACD) in which the signal is registered and counted when the produced high energetic two gamma rays are detected simultaneously (in coincidence) by two detectors opposite each other. Then, the origin of the two photons or the annihilation position is assumed to be somewhere along the line connecting the associated detectors (called the line of response LOR). Time window and energy resolution parameters are used to make sure that the detected gamma rays come from the same annihilation; this in turn helps to accurately localize the spatial distribution of the injected radiolabeled tracer. However, not all photons produced by the annihilation process will reach the detectors as many

of them may undergo absorption or scattering before being detected. Consequently, they are four classes of detected events including true, randoms, scatter and single coincidence events.^{10, 21, 22, 29, 30}

True event occurs when the two high energetic gamma rays come from same annihilation and are successfully detected and counted within time window and energy window, while random event arises when the two high energetic gamma rays do not originate from same annihilation but still they are detected within the time window. To obtain quantitative imaging, subtraction of randoms coincidence events from the total measured counts must be done. Scatter event occurs when one of the detected annihilation photons undergoes a Compton interaction while passing the body, leading to incorrect position information. Attenuation event corresponds to the detection of a single photon within the coincidence time window. The missing photon is due to the high possible interaction of the gamma photons (absorption and scattering) with the human tissues while passing through the body. To obtain quantitative PET information, the proper working of attenuation correction must be assured. This can be done by a transmission scan using external photon point source, or using CT scan.^{10, 21, 22, 29, 30}

1.1.6 Image reconstruction

During data acquisition, the LORs are organized into sets of projections representing the line integrals along all parallel LORs at specific angle. After acquisition and applying the required corrections, in the next step of PET image processing, the measured projections must be converted into an image through process called image reconstruction. This can be done using various algorithms.^{10, 21, 22, 29, 30} The direct and simple one is filtered back projection.²² With the improvement of computing ability, the most commonly used algorithms for PET image

reconstruction are iterative algorithms.²² The results of all reconstruction algorithms and the equality of the reconstructed PET images are influenced by the selection of image matrix sizes and the applied filter. In case of iterative algorithms, an adequate number of iterations and subsets must be chosen for the available statistical quality of the data.³¹

1.1.7 Static and dynamic PET images

PET data is stored in a large matrix and the radioactivity concentration values are assigned to each voxel. Normally they are two types of PET images, namely static and dynamic PET images. In static PET images, we have only one time frame and each voxel contains the summation of the information over the entire image acquisition period. This image type is used when the time dimension is irrelevant. On the other hand, to have more information about different tissues, dynamic PET images with various numbers of frames and different frame durations can be used. With the additional three spatial dimensions, it contains temporal dimension so the distribution of tracer concentration as a function of time can be measured for different metabolic structures inside the human body.^{10, 21, 22, 29, 30, 32}

1.2 PET image segmentation using cluster analysis

Image segmentation is an important task to delineate tumor regions. Different approaches of PET image segmentation have been described, discussed and implemented in literature.³³⁻³⁷

Drawing regions-of-interest (ROIs) directly on PET images is performed by the physicians to identify different functional structures. The tissue time-activity curves (TACs) which describe the

radiotracer concentration as a function of the time can be extracted for further analysis, i.e. to quantify the tumor regions by implementing a compartment model.³⁸ However, the manual delineation of ROIs is a time-consuming, effort-demanding task and is operator dependent. Furthermore, it is challenging due to the inability of the operator to integrate the full TACs of the voxels in the evaluation and of the noise in PET images.^{33, 39, 40} An alternative solution is the automated segmentation of dynamic PET images, which is not only time-saving and highly reproducible, but also takes full time-activity curves (TACs) into consideration.³⁶ There is a great interest on fully-automated approaches with no user interaction to ensure high reproducibility and to reduce the workload of physicians. *k*-means is one of the most frequently used algorithm due to its simplicity.⁴¹⁻⁴⁴

1.2.1 *k*-means clustering

k-means classifies big data into several groups based on a specific criterion in such a way that the objects within a group are similar to each other but different to objects of other groups.⁴¹⁻⁴⁴ For dynamic PET images, it is used to partition voxel TACs into regions with similar kinetic behavior. Each cluster is represented by its centroid which is the average of TACs for all voxels in that cluster. The workflow of the segmentation *k*-means is as follows:

1. The number of the clusters must be defined by the user.
2. The initial cluster centroids are assigned randomly by the algorithm.
3. The distances between each centroid and each voxel-TAC are calculated. The most often used distance measure is Euclidean distance but one can chose alternatives measures such as: city-block, cosine and correlation distance measures.

Therefore, the objective function OF to minimize is the sum of squared distances within each cluster:

$$OF = \sum_{i=1}^{NC} \sum_{j=1}^{n_i} \|\mathbf{x}_j - \mathbf{c}_i\|^2 \quad (1)$$

\mathbf{x}_j is the vector of the TAC of each voxel j , \mathbf{c}_i is the centroid or the mean TAC for cluster i , NC is the number of clusters and n_i is the number of voxels within each cluster.

4. Thus, k -means assigns each voxel-TAC to the closest centroid using the chosen distance measure.
5. When all voxel-TACs are assigned, k -means recalculates the new centroids by taking the average of all voxel-TACs within each individual cluster.
6. This process is repeated until one of the termination conditions (the centroids no longer change or the maximum number of iterations is reached) is fulfilled, otherwise the calculation keeps running.

k -means cluster analysis depends on random centroid initializations; repeated replications may yield different results. Therefore, one may need to repeat the cluster analysis few times to find good clustering solutions and to check the reproducibility of the results. Moreover, the main drawback of the k -means algorithm is that the number of clusters needs to be entered by a user before running the algorithm.^{41, 42} Automated determination of the number of clusters, which affects the results of clustering, is one of the main challenges in data clustering.⁴² To resolve this issue, cluster validity indices are used to validate the goodness of partitions after clustering.

1.2.2 Cluster validity indices

The validation indices are based on the similarity of objects within the same cluster (compactness) and the dissimilarity of objects in different clusters (separation).^{45, 46} Different types of indices exist such as the Bayesian information criterion (BIC), the Akaike information criterion (AIC), Silhouette, Dunn's and WB indices^{47, 48, 49, 50}, which are used to find the optimal number of clusters depending on the dimensionality and noise of the data under consideration.

1.2.3 Validity indices used for PET image segmentation

PET image segmentation is considered as a difficult task because of partial volume effects, bias due to scatter, random, attenuation coincidences, and patient motion during measurement.^{33, 51-53} Consequently, most validity indices have limitations for PET image analysis.⁴⁶ Some of these limitations are related to data characteristics, for example to the noise in the data, or to the dimensionality of the data. In addition, cluster validity indices can also be affected by the sizes, densities and shapes of the clusters.⁴⁶

Different cluster validity indices were proposed as a validation tool for PET image segmentation. For example, for dynamic [¹⁸F]fluoro-deoxy-glucose PET ([¹⁸F]FDG-PET) image segmentation, the Akaike information criterion (AIC) and the Schwarz criterion (SC) were used to define the number of clusters and simulations with the Zubal phantom were used for validation of the results.³⁶ However, testing these indices for dynamic [¹⁸F]FET-PET images segmentation was not performed. In addition, Silhouette and modified Dunn's indices were also used as validation indices when comparing different clustering methods that can be used in the context of dynamic [¹⁸F]FDG-PET data.³⁷ However, which of them could lead to more efficient segmentation of PET images was not investigated.

Accordingly, the Akaike information criterion AIC,⁴⁷ WB,⁵⁰ I ,⁵⁴ modified Dunn's,⁵⁵ and Silhouette⁵⁶ indices were investigated for dynamic [¹⁸F]FET-PET images segmentation based on the developed general requirements for an adequate cluster validity index. AIC was investigated for our datasets as it was used earlier for PET image segmentation.³⁶ WB-index is based on the sum-of-squares within clusters and the sum-of-squares between clusters. It was investigated due to its simplicity and being already used for other multi-dimensional data.⁵⁰ The I -index was investigated by Jegatha Deborah et al.⁵⁷ and it was claimed that this index works well on any type of data sets. This motivated us to use this index for medical image segmentation application. Last, Silhouette and modified Dunn's indices were already used as validation indices by another group in the context of dynamic [¹⁸F]FDG-PET data.³⁷ However, which of them could lead to a better segmentation of dynamic [¹⁸F]FET-PET images was not investigated. Additionally, these cluster validity indices were expected to be valid for the noisy and high dimensional PET data.

1.3 Aims

The objectives of this study were:

- First, to define relevant general requirements for adequate clustering validity index best suited for determining reproducibly and automatically the optimal number of clusters for dynamic data.
- Second, to implement and compare different cluster validity indices with respect to the developed requirements at the example of selecting the best brain image segmentation for [¹⁸F]FET-PET.

1.4 Structure of the thesis

This thesis is divided into five chapters where the first is the introduction including an overview and the state-of-the-art of PET measurements, image segmentation, *k*-means clustering and cluster validity index. The materials and methods used in this study are described in chapter 2. The results and the discussion are presented in chapters 3 and 4, respectively. Chapter 5 gives the conclusions drawn from the study.

2. Materials and Methods

This chapter describes the materials and methods in the order of the work flow. First, the data acquisition, second the data preprocessing, then the clustering, and the estimation of the optimal number of clusters including reproducibility checks. Finally, requirements for an adequate cluster validity index are described and the methods used to check the requirements are documented.

2.1 Patients and data acquisition

Dynamic [^{18}F]FET-PET images were acquired using the PET/CT Biograph mCT (Siemens, Erlangen, Germany) for 45 min in list mode for 8 patients with brain tumors. The data acquisition was obtained directly after [^{18}F]FET tracer injection. Then, the raw data were iteratively reconstructed using the *TrueX* algorithm (Siemens Medical Solutions, Knoxville, TN 37932, USA)³¹ with 24 subsets/4 iterations, 3 mm Gaussian filter and with 21 time frames (6×20 s, 8×60 s, 7×300 s).⁵⁸ To obtain quantitative PET data, the required corrections were performed including dead-time, randoms, attenuation and scatter corrections. The images were sampled into 400×400 voxels with voxel size $2.0 \times 2.0 \times 3.0$ mm³ in x, y and z dimension, respectively and 56 slices per each frame.

2.2 PET data preprocessing

After acquisition of the dynamic PET data, the data underwent two preprocessing steps described in the following.

Mask

To reduce the number of noisy voxels and to extract and include the relevant voxels in the subsequent analysis, a threshold mask was created. The threshold value was set to the mean of all dataset (voxel-activities) of the investigated slice. Thus, only the voxels larger than the threshold were included in the clustering.

Frame weighting

In this study, the data were reconstructed using 21 frames with different frame duration. Hence, the image with shorter duration is noisier (low detected counts) than the image with longer frame duration (high detected counts). To prevent low count frames from dominating the clustering process, a statistical weighting⁵⁹ was applied to the frames j according to:

$$w_j = \frac{(\text{FrameDuration})^2}{P_j + R_j}, \quad (2)$$

with the total counts P_j and the total randoms R_j in frame j .

According to that, the frames with longer duration have higher weighting factors due to higher signal to noise ratio.

2.3 Clustering using the *k*-means algorithm

This algorithm was used to divide voxel TACs according to their shape and magnitude into several groups using the Euclidean distance measure. According to that, the data were converted into two-dimensional matrix, with column corresponding to frame, and row corresponding to voxel and then *k*-means was performed using the function *k*-means of MATLAB (R2015b, MathWorks, Natick, Massachusetts, USA) software. More specifically, this was done using this command: `idx = kmeans(X, k)` to partition the voxel-TACs in the matrix *X* into *k* clusters. Rows of *X* correspond to voxels activity concentration of PET images; columns correspond to the time frames. *k*-means returns a vector (*idx*) containing the cluster indices of each voxel-TAC. In this study, the maximum number of iterations was set to one million which was never reached.

2.4 Reproducibility of *k*-means objective function value

To measure the reproducibility of *k*-means, the coefficient of variation (*CV*), i.e. the ratio of standard deviation to mean, was calculated for the objective function values *OFVs*. The calculation was done using 100 random centroid initialization replications (*RCI*₁₀₀), for number of clusters ranging from 2 to 50. Maximum *CV* was given as indicator of *k*-means reproducibility.

2.5 Cluster validity indices

To determine the optimal number of clusters for dynamic brain [^{18}F]FET-PET images, the following indices were investigated (TABLE I):

- AIC^{47} is a criterion used for model selection and it is a function of the sample size, the number of clusters (NC) and the sum of squares distances within clusters (SSW). The optimal number of clusters is assumed to be the value for which the AIC is minimal.
- WB-index^{50} is a function of NC (the number of clusters), SSW (sum-of-squares within cluster) and SSB (sum-of-squares between clusters). The optimal number of clusters is assumed to be the value for which the WB-index is minimal.
- $I\text{-index}^{54}$ measures the separation based on the maximum Euclidian distance between cluster centroids and measures compactness based on the sum of distances between voxel-TACs and their cluster center. The optimal number of clusters is defined where this index reaches its maximum.
- The modified Dunn's index^{55, 60} computes the distance between the voxel-TACs within a cluster to the centroids of different clusters as the inter-cluster separation and the maximum diameter among all clusters as the intra-cluster compactness. The optimal cluster number is defined where this index reaches its maximum.
- The Silhouette index⁵⁶ is based on the pairwise difference of between and within cluster distances to validate the clustering performance. The optimal number of clusters is defined where this index reaches its maximum.

TABLE I. Cluster validity indices.

Measure	Definition	Optimal value
AIC	$AIC = N \times \ln \left(\frac{\sum_{i=1}^{NC} \sum_{j=1}^{n_i} \ x_j - c_i\ ^2}{N} \right) + 2(NC + 1)$	Min
WB-index	$WB = NC \cdot \frac{SSW}{SSB}$ $Compactness = SSW = \sum_{i=1}^{NC} \sum_{j=1}^{n_i} \ x_j - c_i\ ^2$ $Separation = SSB = \sum_{i=1}^{NC} n_i \ c_i - \bar{X}\ ^2$	Min
I-index	$\left(\frac{1}{NC} \cdot \frac{\sum_{n=1}^N \ x_n - \bar{X}\ }{Compactness} \cdot Separation \right)^p$ $Separation = \max_{i,j} \ c_j - c_i\ $ $Compactness = \sum_{i=1}^{NC} \sum_{j=1}^{n_i} \ x_j - c_i\ $	Max

Modified

$$\min_i \left(\min_j \left(\frac{Separation_{i,j}}{Compactness} \right) \right)$$

Max

Dunn's

index

$$Separation_{i,j} = \frac{1}{n_i + n_j} \left(\sum_{i=1}^{n_i} \|x_i - c_j\| + \sum_{j=1}^{n_j} \|x_j - c_i\| \right)$$

$$Compactness = \max_k \left(\max_{x,y \in C_k} \|x - y\| \right)$$

Silhouette

$$\frac{1}{NC} \sum_i \left(\frac{1}{n_i} \sum_{x \in C_i} \frac{Separation_i - Compactness_i}{\max[Separation_i, Compactness_i]} \right)$$

Max

index

$$Separation_i = b(x)$$

$$Compactness_i = a(x)$$

D: data set; x_i : TAC of each voxel i ; N: number of objects in D; $\bar{\mathbf{X}} = \frac{\sum_{i=1}^N x_i}{N}$: center of D; NC: number of clusters; p: number of attributes in D; n_i : number of objects in C_i ; C_i : i^{th} cluster; c_i : center of C_i ; $d(\mathbf{x}, \mathbf{y})$: distance between \mathbf{x} and \mathbf{y} ; $a(\mathbf{x})$: average distance between the object \mathbf{x} to the other objects in the same cluster as \mathbf{x} ; $b(\mathbf{x})$: minimum average distance between the object \mathbf{x} to objects in a different cluster, minimized over all clusters; SSW: sum of square within cluster; SSB: sum of square between clusters.

2.6 Determination of the needed number of replications for each validity index

k -means cluster analysis depends on random centroid initializations, repeated replications may yield different results. Therefore, one needs to repeat the cluster analysis a few times to find good clustering solutions and to check the reproducibility of the results. The required number of replications needed by each validity index was therefore investigated. This was done in 2 steps:

First step: The true optimal number of clusters was calculated using 100 random centroid replications for each number of clusters ranging from 2 to 50 (RCI_{100}). The replication with the minimal objective function value was selected for each cluster number from the 100 centroid initialization replications. Then the optimal number of clusters was determined by the global extremum of WB, I , modified Dunn's and Silhouette validity indices.

Second step: A single random centroid initialization for each number of clusters ranging from 2 to 50 was performed (RCI_1) and the optimal number of clusters suggested by each validity index was collected. This process was replicated 100 times (so in total we have 100 numbers of clusters suggested by each index for each investigated slice for each patient ($100 \times RCI_1$)). The optimal number of clusters for each single random centroid initialization (RCI_1) was compared with the true optimal number of clusters determined in step one. The probability of getting a wrong result was estimated from the number of wrong results divided by 100. Then the probability to obtain a wrong optimal cluster number for n runs was estimated to be $(p_{false})^n$. Thus, n was estimated from:

$$(p_{false})^n \leq threshold, \text{ with } threshold = 1.0 \times 10^{-7}. \quad (3)$$

This procedure was performed 24 times (for the 3 neighboring slices for all 8 patients).

2.7 Evaluation of the validity indices for the general requirements

- **Requirement number 1 (the number of clusters suggested by the index should give a physiologically reasonable segmentation)**

To check the reasonability of cluster result, visual inspection was performed on the images of clusters and their TACs. At least 3 clusters (white/gray matter, and tumor) were expected and any other result was not accepted.

- **Requirement number 2 (the number of clusters suggested by the index should not depend on the number of the investigated voxels)**

To test the dependence of the optimal number of clusters according to the validity indices on the number of voxels, k -means clustering analysis was applied after duplication of a slice selected from each patient. The result was compared with applying cluster analysis for same slice without duplication.

- **Requirement number 3 (the optimal number of clusters suggested by the index should be highly reproducible)**

In addition to investigating the reproducibility of k -means, the reproducibility of the tested indices was also evaluated and compared. Thus, the dependence of each validity index on the randomly initialized centroid replications was investigated as follows: For each number of clusters (ranging from 2 to 50), 100 randomly initialized centroid replications of k -means were calculated. From the results 100 values for each validity index and number of clusters were obtained. The coefficient of variation (the ratio of standard deviation to mean) was calculated for each index for every cluster number and for every patient. As global indicator for the dependence

of a validity index on the randomly initialized centroid, the maximum coefficients of variation of the patient population were calculated at the optimal number of clusters. Furthermore, to guarantee highly reproducible results, the required number of replications was calculated for WB, I , modified Dunn's, and Silhouette indices using Eq. (3) for further demonstration of the dependence of each index on random centroid initializations.

- **Requirement number 4 (the index should have a clear single extremum)**

To verify whether the validity indices have a single extremum, visual inspection was performed on the replication with minimum OFV from 100 random centroid initialization replications (RCI_{100}) for number of clusters ranging from 2 to 50.

- **Requirement number 5 (the computation time should be reasonable for use in a clinical setting)**

Indices calculation time needed for single replication was recorded for one slice (119×97 voxels \times 21 frames) of one patient with number of clusters fixed to 10.

3. Results

In this chapter, the results of the reproducibility of k -means and the investigating of the cluster validity indices with respect to the developed requirements of an adequate cluster validity index are presented.

FIG.1 and FIG.2 show a typical slice each after reconstruction and before clustering in 21 frames. These patients had oligodendroglioma and glioblastoma with corresponding high tracer uptake, respectively.

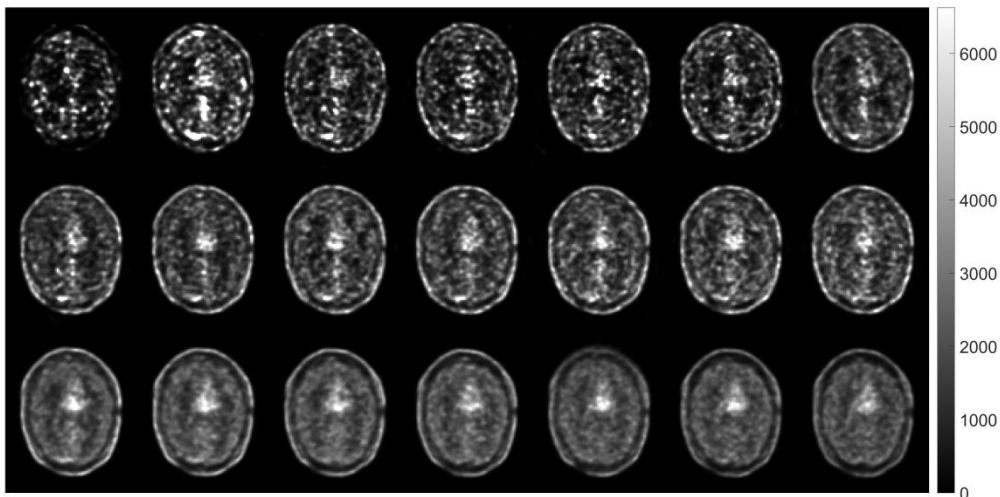


FIG.1. Reconstructed dynamic $[^{18}\text{F}]\text{FET}$ -PET images for a patient had oligodendroglioma. Note that concentrations higher than 6000 (Bq/cc) were set to 6000 (Bq/cc).

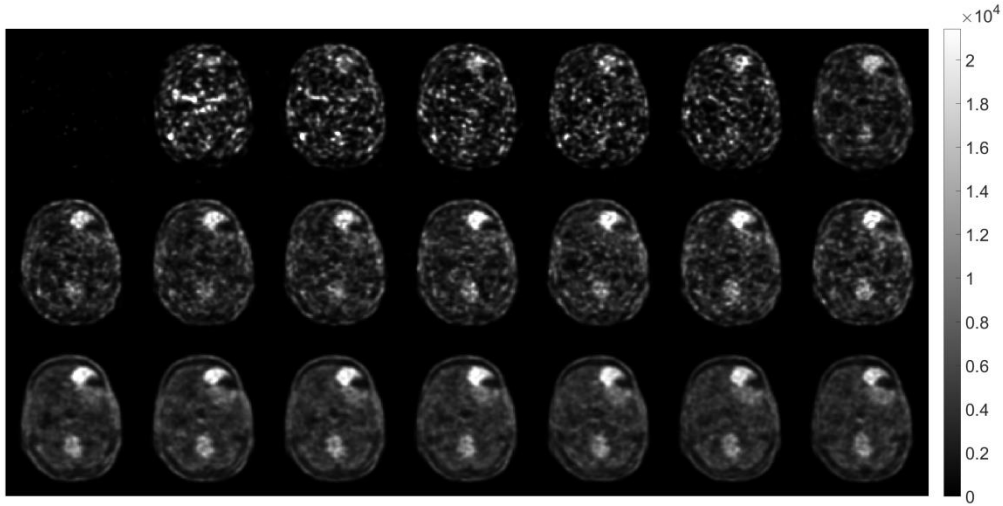


FIG.2. Reconstructed dynamic $[^{18}\text{F}]\text{FET-PET}$ images for a patient had glioblastoma. Note that concentrations higher than 21415 (Bq/cc) were set to 21415 (Bq/cc).

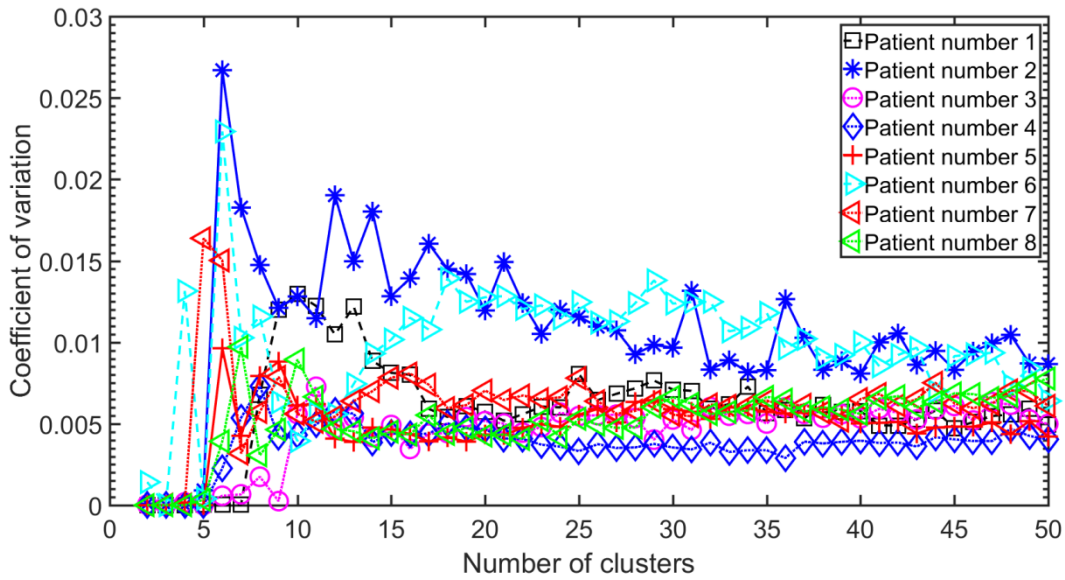


FIG.3. Coefficient of variation (CV) of the k -means objective function value (OFV) using 100 random centroid initialization replications (RCI_{100}) to evaluate the reproducibility of k -means, for number of clusters from 2 to 50.

3.1 Reproducibility of k -means objective function value

Within all patients, the maximum CV of OFV was 2.7×10^{-2} as depicted in FIG.3. The CV strongly depends on the patient and on the number of clusters.

3.2 Cluster validity index requirements

TABLE II summarizes the developed general requirements during this study, which were used as a basis for the evaluation. Accordingly, to evaluate the optimal number of clusters, the validity indices were investigated according to the five requirements (TABLE II). The first requirement (physiologically reasonable segmentation) is shown at the end of this section as the optimal number of clusters must be defined first using the different validity indices before evaluating the tissue segmentation result.

TABLE II. General requirements for the optimal number of clusters determined based on an adequate cluster validity index.

Requirement	Specification
number	
1	The optimal number of clusters should give a physiologically reasonable segmentation.
2	The number of clusters should not depend on the number of voxels included in the analysis.*
3	The suggested optimal number of clusters should be highly reproducible.£
4	The index should have a clear single extremum to specify the optimal number of clusters uniquely.
5	The required computation time should be reasonable for use in a clinical setting.

*For example doubling each voxel should yield the same optimal number of clusters.

£ For example the difference of the validity index value between clustering results with different numbers of clusters should be large compared to those within different replications of a given number of clusters.

Requirement number 2

Based on the AIC, different optimal numbers of clusters were suggested when applying cluster analysis to duplication of a slice and the same slice without duplication. This demonstrates that this index does not fulfil the needed second requirement for an adequate validity index (TABLE II). The dependence of this index on the number of voxels can be seen in FIG.4. AIC performance thus was inadequate and was excluded from further comparison.

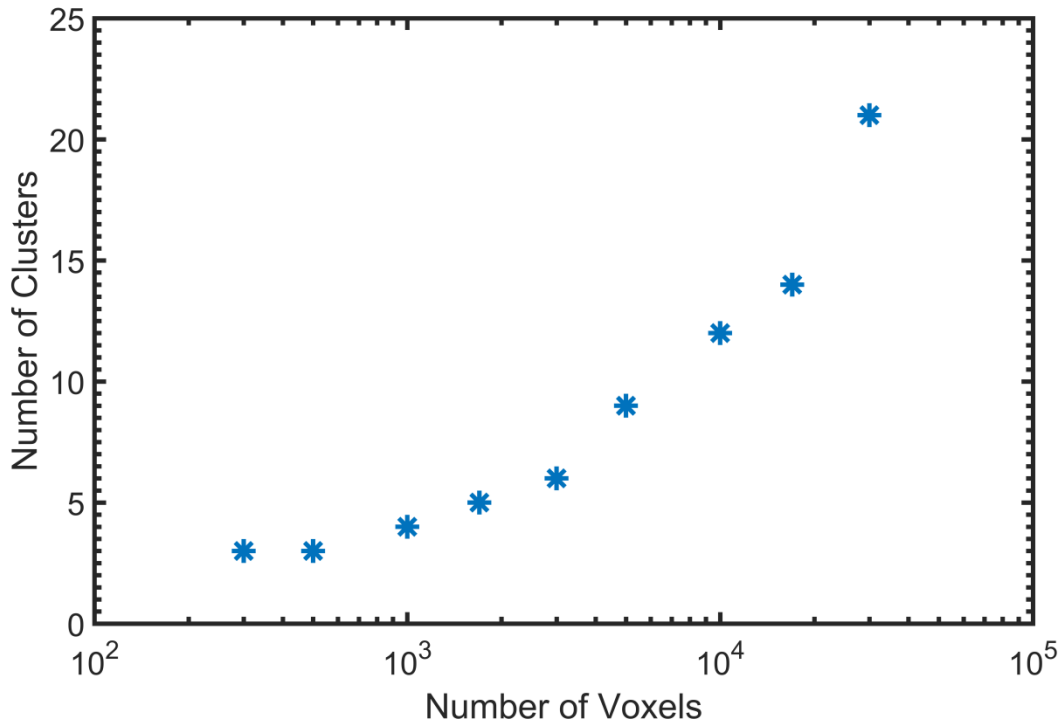


FIG.4. Optimal number of clusters suggested by the AIC as a function of the number of voxels (range 300 – 30000) showing the violation of requirement 2 (TABLE II; note the logarithmic scale for the x-axis).

At the same time, the identical optimal number of clusters was observed between applying cluster analysis to duplication of a slice and the same slice without duplication for WB, I , modified Dunn's, and Silhouette indices in all patients, suggesting that all these four validity indices meet and fulfil this requirement.

Requirement number 3

The dependence of the four indices on the random centroid initialization for each number of clusters (ranging from 2 to 50) for the patients shown in FIG.1 and FIG.2 is seen in FIG.5 and FIG.6, respectively. There are deviations among replications within a given number of clusters. I and modified Dunn's indices have lower reproducibility and are strongly influenced by the initial centroids selection. This influence of random centroid initialization is less prominent in the other indices. From all patients, the maximum coefficient of variation of WB, I , modified Dunn's, and Silhouette validity indices were 3×10^{-2} (optimal number of clusters 6), 1 (optimal number of clusters 6), 2×10^{-1} (optimal number of clusters 3) and 3×10^{-3} (optimal number of clusters 2), respectively.

The required numbers of replication needed to yield with high reproducibility the optimal number of clusters according to Eq. (3) were 4, 39, 41, and 23 for WB, I , modified Dunn's, and Silhouette validity indices, respectively.

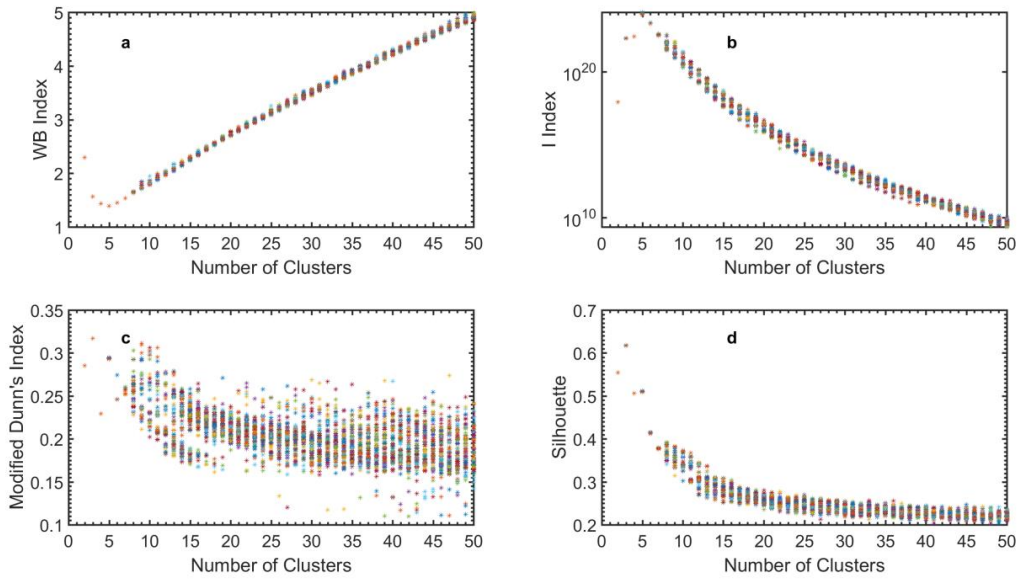


FIG.5. Cluster validity indices plot for the number of clusters from 2 to 50 with 100 different random initial centroids of the patient shown in FIG.1. At the optimal number of clusters, defined by WB, I (note the logarithmic scale for the y-axis), modified Dunn's and Silhouette indices, the coefficients of variation were 7×10^{-5} , 2×10^{-1} , 2×10^{-4} and 2×10^{-4} , respectively.

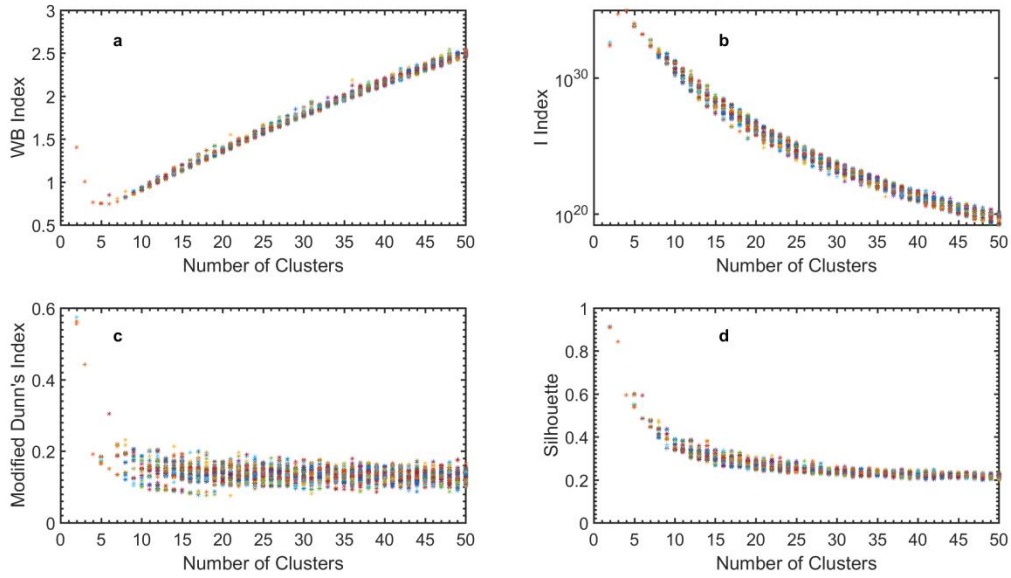


FIG.6. Cluster validity indices plot for the number of clusters from 2 to 50 with 100 different random initial centroids of the patient shown in FIG.2. At the optimal number of clusters, defined by WB, I (note the logarithmic scale for the y-axis), modified Dunn's and Silhouette indices, the coefficients of variation were 3×10^{-2} , 8×10^{-16} , 2×10^{-2} and 3×10^{-3} , respectively.

Requirement number 4

To determine the optimal number of clusters, the replication with the minimum objective function value was selected from 100 centroid initialization replications (FIG.5 and FIG.6). To have a better visualization, we showed the data with logarithmic scale for the y-axis of I -index. The optimal number of clusters was determined by the global extremum of WB, I , modified Dunn's and Silhouette validity indices (FIG.7 and FIG.8). Within all patients, only the WB-index showed a single global maximum, whereas the other indices showed also local extrema.

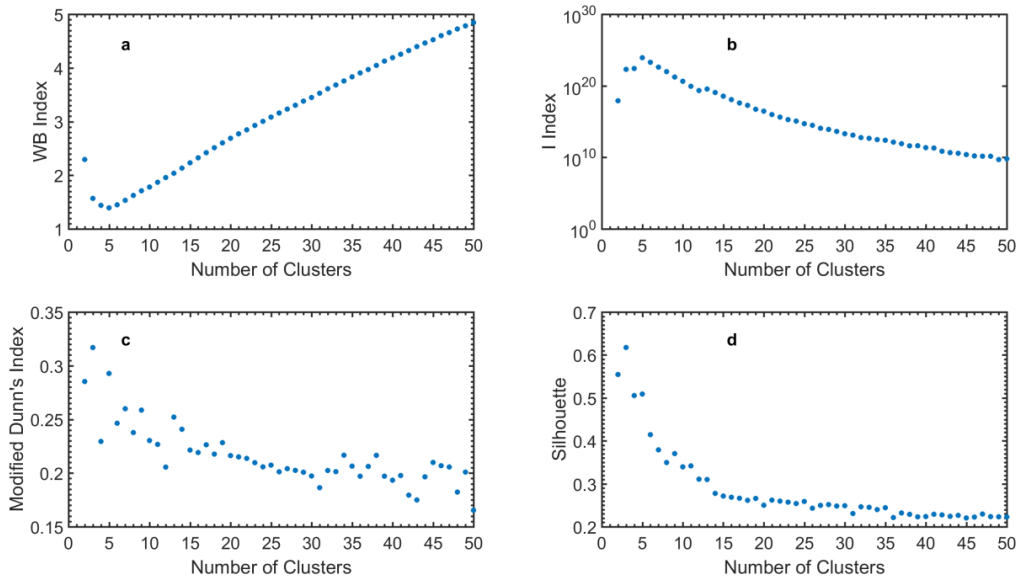


FIG.7. The optimal number of clusters is defined for a given range from 2 to 50 according to (a) WB, (b) I (note the logarithmic scale for the y -axis), (c) modified Dunn's and (d) Silhouette validity indices for the patient shown in FIG.1. The replication with the minimal objective function value was selected from 100 replications.

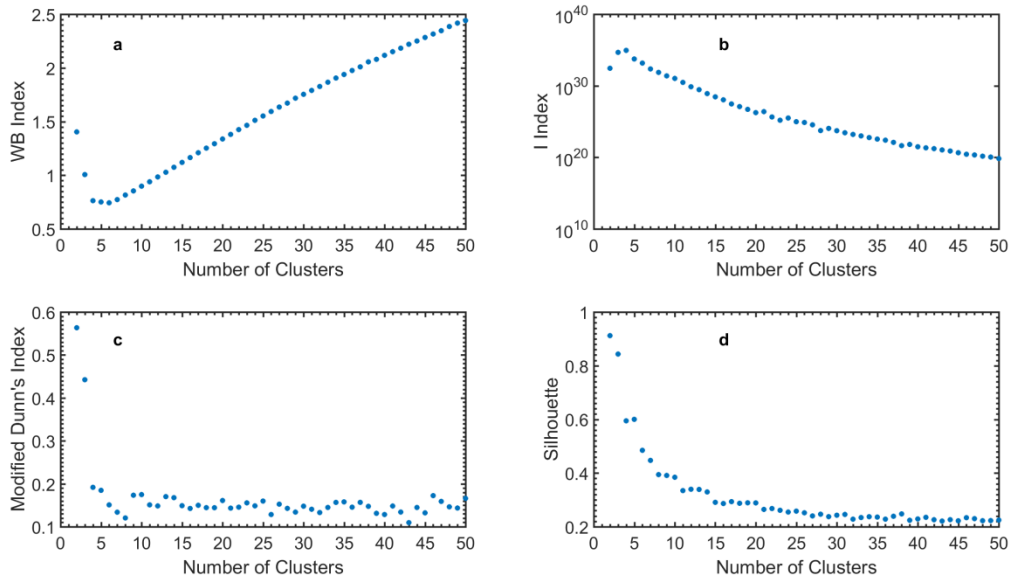


FIG.8. The optimal number of clusters is defined for a given range from 2 to 50 according to (a) WB, (b) I (note the logarithmic scale for the y -axis), (c) modified Dunn's and (d) Silhouette validity indices for the patient shown in FIG.2. The replication with the minimal objective function value was selected from 100 replications.

Requirement number 5

The calculation times needed (in seconds) using one replication were 0.2 ± 0.1 , 0.0004 ± 0.0002 , 0.05 ± 0.02 , 0.0009 ± 0.0002 , and 2.5 ± 0.1 for k -means, WB-index, I -index, modified Dunn's index and Silhouette index, respectively. Thus, computation time is not an issue when using validity indices.

Requirement number 1

k -means cluster segmentation results for the patients shown in FIG.1 and FIG.2 are presented in FIG.9 and FIG.10, respectively for the optimal number of clusters suggested by the WB, I , modified Dunn's, Silhouette validity indices. White matter, gray matter, and 2 regions within the tumor were clearly delineated with the optimal number of clusters suggested by WB and I indices, compared with the number of clusters suggested by modified Dunn's and Silhouette indices (FIG.9). In all patients, the tumor was separated from the whole brain with the optimal number of clusters suggested by modified Dunn's, Silhouette validity indices.

The extracted tumor TACs obtained using the corresponding optimal number of clusters suggested by WB and Silhouette indices for the patients shown in FIG.1 and FIG.2 are depicted in FIG.11 and FIG.12, respectively. Additional sub-regions inside the tumor are observed by cluster analysis using WB-index with noisy TAC due to the lower number of voxels (110) within that cluster shown in FIG.11.

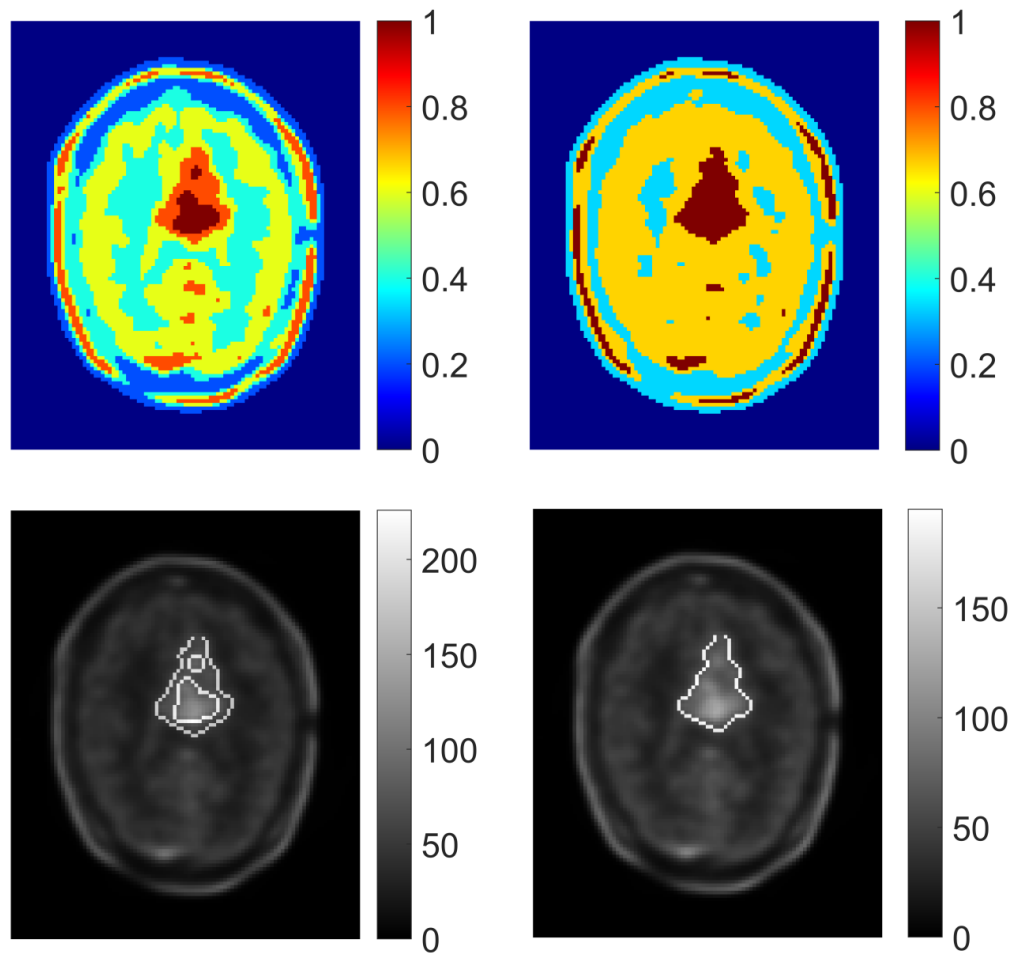


FIG.9. k -means segmentation (upper row) and superimposed tumor contours (lower row) of the $[^{18}\text{F}]$ FET-PET kinetics of the patient with an oligodendroglioma (FIG.1) for the optimal number of clusters $NC=5$ given by WB and I indices (left column) and the optimal number of clusters $NC=3$ given by the modified Dunn's and Silhouette indices (right column). Different clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0. The contour was detected using the 'log' edge algorithm (Laplacian operator) of MATLAB (R2015b, MathWorks, Natick, Massachusetts, USA) software. The obtained contours were checked by an experienced physician for plausibility. Non-tumor cluster areas such as veins, arteries and the sinus cavernous or enrichment of the mucous membranes of the nasal sinuses in the $[^{18}\text{F}]$ FET-PET brain images are not contoured.

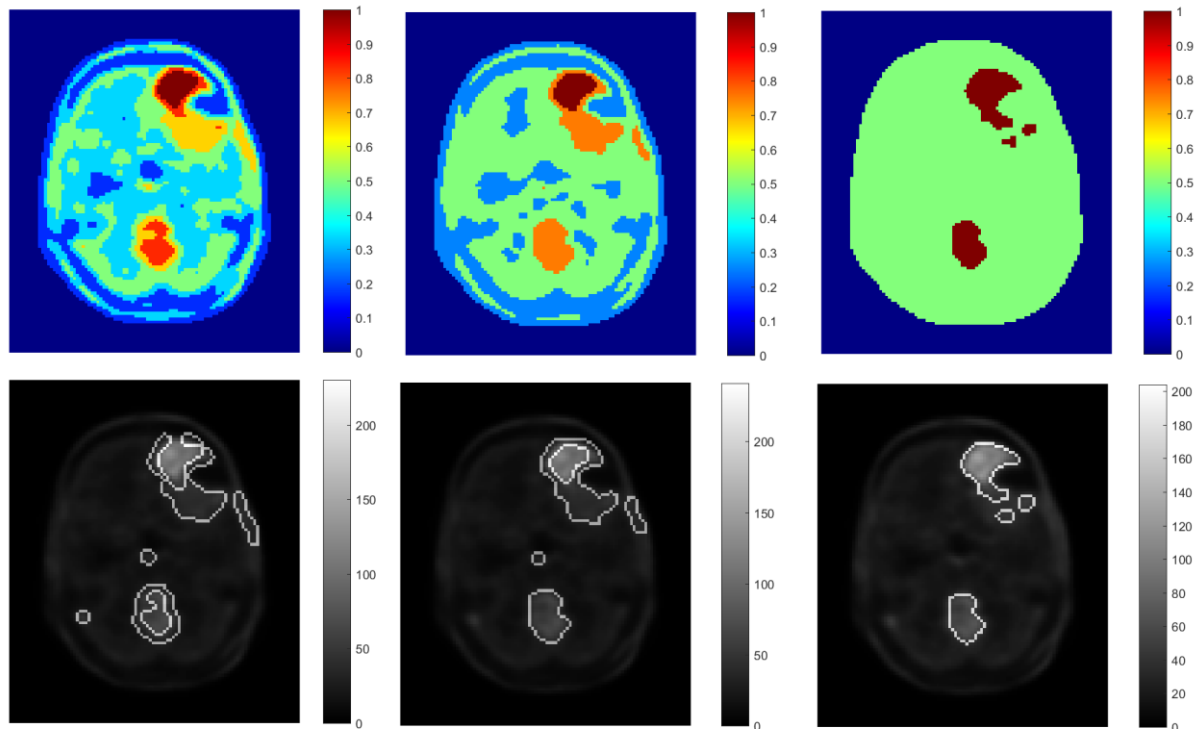


FIG.10. k -means segmentation (upper row) and superimposed tumor contours (lower row) of the $[^{18}\text{F}]\text{FET-PET}$ kinetics of the patient with glioblastoma (FIG.2) for the optimal number of clusters $NC=6$ given by WB-index (left column) and the optimal number of clusters $NC=4$ given by I index (middle column) and the optimal number of clusters $NC=2$ given by the modified Dunn's and Silhouette indices (right column). Different clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0.

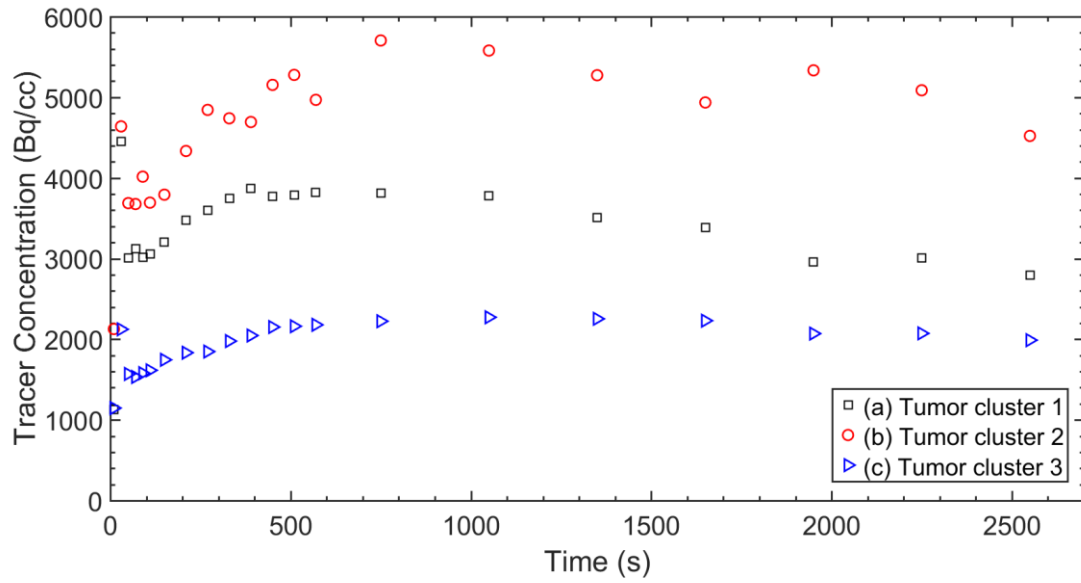


FIG.11. TACs corresponding to the tumor clusters (seen in FIG.9) based on the optimal number of clusters suggested by the WB-index or *I*-index (a) and (b) and by the optimal number of clusters suggested by modified Dunn's index or Silhouette index (c). The numbers of voxels were 597, 110, and 601, for tumor clusters 1, 2 and 3, respectively.

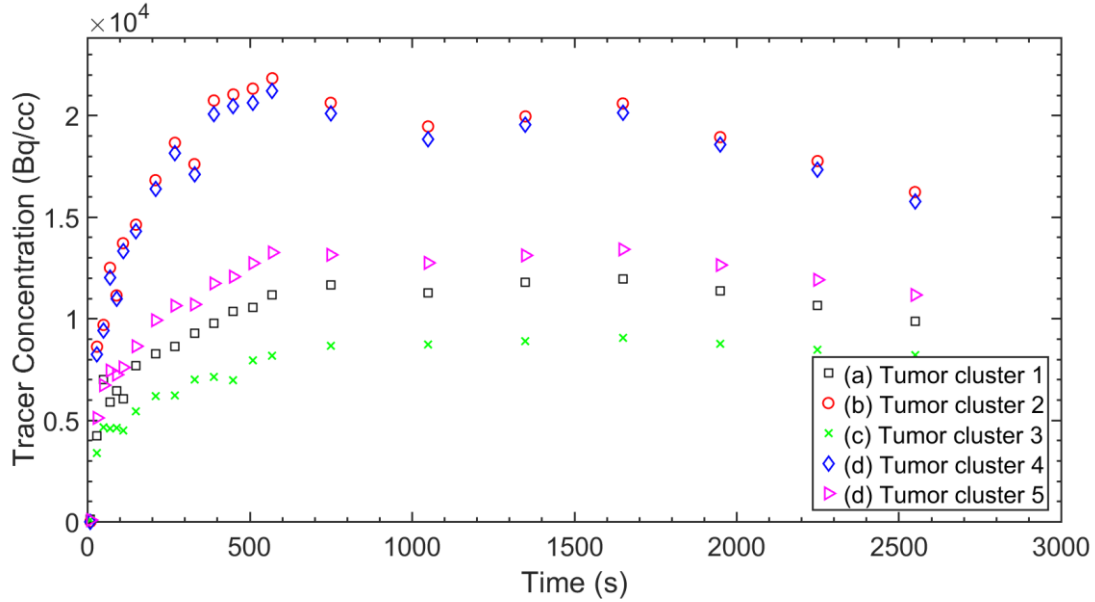


FIG.12. TACs corresponding to the tumor clusters (seen in FIG.10) based on the optimal number of clusters suggested by the WB-index (a) and (b), the optimal number of clusters suggested by *I*-index (c) and (d) and the optimal number of clusters suggested by modified Dunn's index or Silhouette index (d) and (f). The numbers of voxels were 175, 115, 498, 128 and 414 for tumor clusters 1, 2, 3, 4 and 5, respectively.

The clustering algorithm was performed for every patient independently on 3 neighboring slices containing tumor tissue. Single slices were investigated because of computational reasons. As only validity indices were included in the evaluation, which do not depend on the number of voxels, the number of optimal clusters should not depend on the number of clustered voxels. Note that a 3D data set has just more voxels from the view of the *k*-means algorithm. The separate clustering of 3 neighboring slices was done to investigate the stability of the optimal number of clusters as the slice thickness is 3 mm. The similarity of the optimal number of clusters was

interpreted as stability based on the assumption that neighboring slices should have very similar if not identical optimal number of clusters.

When considering reasonability of the number of clusters produced by indices, at least three clusters including gray matter, white matter and tumor were expected. The optimal number of clusters given by modified Dunn's and Silhouette indices led to a very poor segmentation (TABLE III) as mostly 2 regions were identified (tumor and remainder). WB and *I* indices suggested higher number of clusters than modified Dunn's and Silhouette indices. Only in two cases *I*-index suggested a higher value, in 14 cases they are the same and in 8 cases WB-index suggested a higher value. The optimal number of clusters ranged from 4 to 6 by WB-index and 3 to 6 by *I*-index. Both indices allowed better differentiation of different tissues as compared with modified Dunn's and Silhouette validity indices. The optimal numbers of clusters suggested by WB and *I* indices deviate maximum by two clusters between neighboring slices.

Performing *k*-means on the whole brain volume (56 slices) is time consuming (6 h per one replication). The sensitivity of the PET is higher in the center of the axial field of view than at the edge of the axial field of view. This results in different statistical properties of the slices depending on the axial field of view. Thus, clustering of three slices by three slices could be more adequate than clustering of all 56 slices together. Therefore, to check the influence of a larger number of voxels on the result, the clustering was performed on the same three neighboring slices together. The optimal numbers of clusters given by the investigated validity indices are listed in TABLE III (numbers in the brackets). Mostly, WB and *I* indices suggested for three neighboring slices together the rounded mean of the numbers of the three neighboring slices separately. Only in two patients the clustering of the 3 slices together yielded a different result between WB and *I* indices. A factor 3 increase of voxels did not largely change the found number of optimal

clusters. Connectivity of the tumor clusters was checked: In all patients the tumor voxels in the 3 slices were spatially connected.

TABLE III. Comparison of the optimal number of the clusters for three neighboring slices as defined by different cluster validity indices. Note: From first row to the eighth row, the first three numbers in each slot are the optimal number of clusters corresponding to three neighboring slices separately. The number in the bracket is the optimal number of clusters corresponding to the same three slices together.

Patient number	WB-index	<i>I</i>-index	Modified Dunn's index	Silhouette index
1	4, 5, 5 (5)	4, 5, 5 (5)	3, 3, 3 (3)	3, 3, 3 (3)
2	5, 6, 5 (5)	4, 4, 4 (4)	2, 2, 2 (2)	2, 2, 2 (2)
3	5, 6, 6 (6)	6, 6, 5 (6)	2, 2, 2 (2)	2, 2, 2 (2)
4	4, 4, 4 (4)	4, 4, 4 (4)	2, 2, 2 (2)	2, 2, 2 (2)
5	5, 4, 4 (4)	4, 4, 3 (4)	2, 2, 2 (2)	2, 2, 2 (2)
6	5, 5, 5 (5)	4, 5, 5 (5)	3, 3, 3 (3)	3, 3, 3 (3)
7	4, 4, 4 (4)	4, 4, 4 (4)	3, 3, 3 (3)	3, 3, 2 (3)
8	6, 5, 4 (4)	5, 6, 4 (7)	2, 2, 2 (2)	2, 2, 2 (2)
Min	4 (4)	3 (4)	2 (2)	2 (2)
Max	6 (6)	6 (7)	3 (3)	3 (3)

An example of the k -means cluster segmentation results of three neighboring slices independently are visualized in FIG.13 and FIG.14, according to the optimal number of clusters suggested by WB validity index.

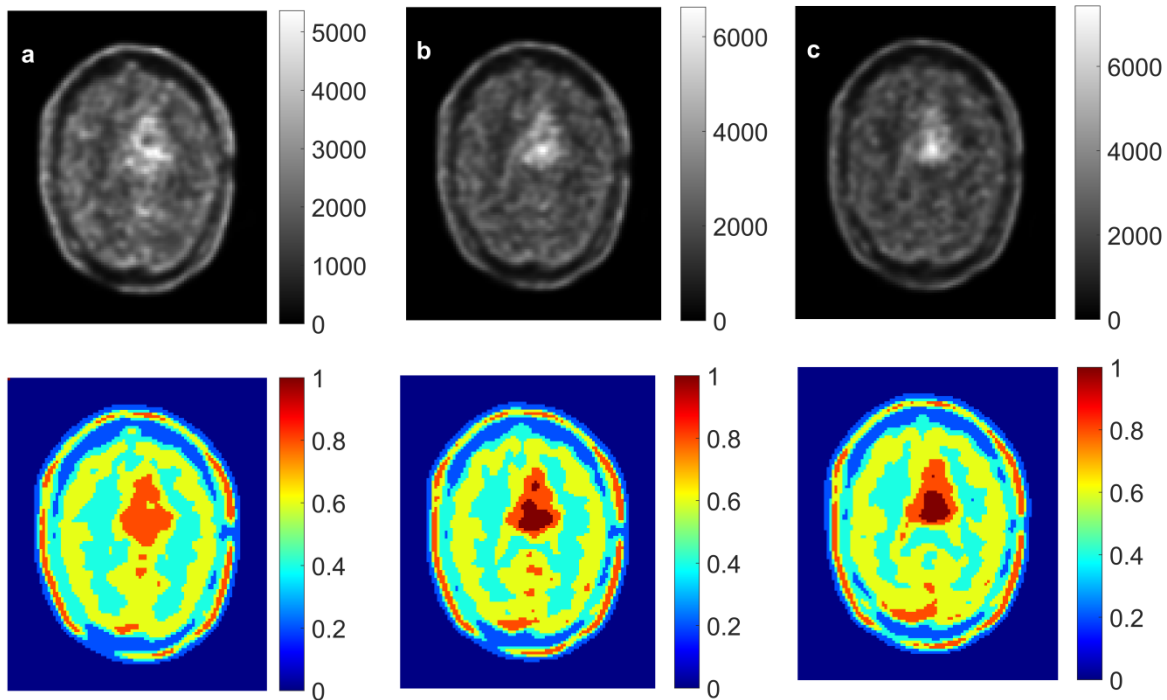


FIG.13. The top row shows an example for three neighboring slices having tumor in the last frame, (a) slice 25 (b) slice 26 (c) slice 27 taken from patient shown in (FIG.1). k -means segmentation result can be seen in the bottom row for the three neighboring slices with the optimal number of clusters, (a) 4 (b) 5 (c) 5, given by WB validity index. Clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0. The slices **b** and **c** have one cluster more than slice **a** due to heterogeneity of tumor in the selected plane.

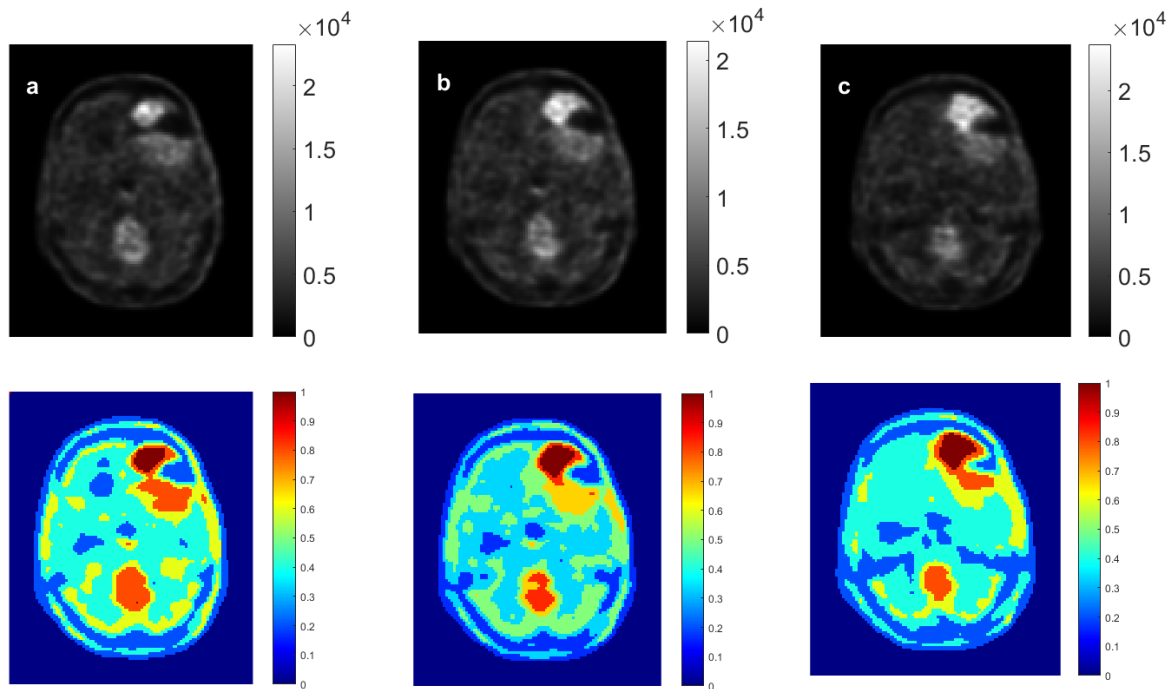


FIG.14. The top row shows an example for three neighboring slices having tumor in the last frame, (a) slice 42 (b) slice 43 (c) slice 44 taken from patient shown in (FIG.2). *k*-means segmentation result can be seen in the bottom row for the three neighboring slices with the optimal number of clusters, (a) 5 (b) 6 (c) 5, given by WB index. Clusters are represented by different colors. The area removed by thresholding corresponds to the dark blue color with value 0. The slices **b** has one additional cluster more than slices **a** and **c**.

4. Discussion

Automated cluster analysis of PET image segmentation is used to accurately and reproducibly define the spatial distribution of the tumor regions. In this study a validation strategy of k -means clustering algorithm for dynamic human brain [^{18}F]FET-PET images is provided. Eight patient data reconstructed for the same parameters on the same scanner were used.

The traditional solution to find the optimal number of clusters is based on cluster validity indices; these however are affected by data characteristics, like noise or the dimensionality of the data. General requirements for a validity index to be adequate to describe PET data are that the number of clusters suggested by the index should result in a physiologically reasonable segmentation, not depend on the number of the investigated voxels and be highly reproducible. Additionally, the index should have a clear single extremum to have a unique solution and the computation time should be reasonable for use in a clinical setting (TABLE II). AIC, WB, I , modified Dunn's and Silhouette validity indices were compared to determine the optimal number of clusters that best fits the given PET data.

Comparison of the five investigated validity indices based on the predefined requirements

AIC was excluded from further comparison as the optimal number of clusters suggested by this index depends on the number of investigated voxels, which is not desired (FIG. 4). Besides, Silhouette and modified Dunn's indices performance was inadequate in our comparison due to different reasons. For example, Silhouette index uses the average minimum distance between clusters as the indicator of the inter-cluster separation, which should be maximized. For a data set with sub-clusters, the inter-cluster separation will achieve higher values when sub-clusters are considered as one big cluster.⁴⁶ Therefore, in our data we cannot observe the sub-cluster within the tumor and other relevant clusters in the brain (FIG.9 (right column)). The main disadvantages of the modified Dunn's index are that on the one hand this index can be potentially influenced by the outlier in the data (as the maximum cluster diameter can be large in a noisy data set).⁴⁶ On the other hand, this index is strongly influenced by the initial centroids selection. Consequently, Silhouette and modified Dunn's indices seem not to be qualified for segmentation of dynamic human brain [¹⁸F]FET-PET images. Furthermore, both indices suggested less number of clusters leading to poor segmentation compared to the number of clusters suggested by WB and *I* indices.

When comparing *I*-index with WB-index, both indices provided similar results. In more detail, WB-index suggested sometimes one or two more clusters than *I*-index and in most cases both suggested the same optimal number of clusters. However, WB-index is more resistant to noise because it takes mean values of all data into account through dividing sum of within-cluster-distances by sum of between-cluster-distances. As compared with *I*-index, WB-index is also simpler to understand. The WB-index showed only one global minimum for the given number of clusters ranging from 2 to 50 for all investigated patients while the other indices did not hold this property.

Within 8 patients, different optimal numbers of clusters were expected depending on the heterogeneity of the tumor and the location of slices. Assuming similar situation between neighboring slices of a specific patient, the optimal numbers of clusters are more or less the same among these slices. In TABLE III, the intra-patient difference of the number of clusters was less than or equal to two clusters between neighboring slices, given by WB and I indices.

Comparison of the manual and the automated segmentation methods

The tumor cluster results obtained were checked by a physician for plausibility. In no case it was deemed that the clustering of the tumor was wrong. As we are confident that the automatic segmentation is more accurate than the manual segmentation, which uses static images (i.e. less information) and is operator dependent, time-consuming and challenging due to noise in the PET images, we do not expect relevant information from comparing the cluster results to the inferior manual method.⁵⁹ In addition, we do not expect relevant information from comparing the cluster results with the manual delineation of ROIs on dynamic images as it is difficult (if not impossible) for the physician to integrate the full voxel time activity curves (i.e. 21 frames with different frame duration and noise properties) into the evaluation (FIG.1 and FIG.2).

Comparison between global k -means algorithm and the current used approach

To further demonstrate the validity of our current used approach, one of the deterministic clustering approaches (the global k -means algorithm)⁶¹ was implemented, investigated and the results were compared to the used repeated random initialization approach. Regarding to the objective function value, we observed that the k -means objective function value of the used approach outperformed the deterministic clustering global k -means in 67 of the 72 investigated cases (93 %). This was calculated for the number of clusters from 2 to 10 (i.e. 9) based on one slice containing parts of tumor for eight patients. Moreover, the optimal numbers of clusters

suggested by WB-index, modified Dunn's index and Silhouette index were the same using both approaches while for the *I*-index, only 1 case out of 24 the optimal number of clusters showed a difference of one in the optimal number of clusters. However, this was to be expected as the *I*-index is sensitive to the random initial centroid selection (in FIG.5 and FIG.6).

Real ground truth data for evaluating dynamic PET image segmentation

Brain PET image segmentation requires a real ground truth which reflects the characteristics of the segmentation problem encountered in practice. Physical and digital phantoms have an important role in assessment and validation of the accuracy of automated segmentation.

To deduce valid conclusions on the correct segmentation in the real patient from a simplified physical phantom is however virtually impossible. This is due to the unknown tumor structure (e.g. perfusion, distribution of different cell types, vessels, etc.) and the tumors are most often simulated as spheres with uniform activity.^{62, 63} In addition, such phantoms typically can neither reproduce the complexity of the brain and nor both the normal and pathological anatomical variability observed in clinical data.⁶⁴

Therefore, using digital phantom as surrogate for real PET data is a common approach to know the ground truth.^{33, 63, 65-70} However, this requires constructing phantoms that reproduce the full range of PET imaging characteristics encountered in practice coming from data acquisition using complex scanner geometry and a reconstruction algorithm. These characteristics include partial volume artifacts (e.g. different percentages of blood vessels in each voxel), poor spatial resolution and noise; these make the construction of adequate digital phantoms a difficult task.⁶⁴ In addition, one needs to generate phantoms with various anatomies and complex shapes and different levels of activity in the tumors in order to generate realistic tumors.^{62, 63, 70}

Moreover, assigning TACs extracted from PET patients to a digital phantom will turn the real problem into an artificially simplified (i.e. one needs to choose the number of TACs used) task,

the relation of which to the real segmentation task is unknown (as the PET data clearly are more complex and do not consist of only a few different TACs). Thus, the results obtained from the simulations are invalidated; this is especially significant when choosing a low number of TACs. In case a very high number (e.g. 1000) of TACs is chosen for the simulation to be nearer to the real situation, then, after adding the noise, a low number of clusters is chosen by all the validity indices; thus, one cannot decide which validity index is better (maybe except for assuming that the best one is the one choosing the highest number of clusters). As additional drawback associated with using the patient images as an input for simulations is that the patient images suffer from limited spatial resolution and noise. The limited spatial resolution and noise propagate through the simulation and reconstruction processes, which result in simulated images different from those initially used as input images.⁷¹ For example, the modulation transfer function (MTF) is applied twice to the data: first during the real measurement, and second during the generation of the sinogram based on the real measurement followed by the image reconstruction.

Using a realistic simulation therefore needs a long computation time and is highly demanding in computational resources.¹⁹ In addition, the existence of simulated datasets is very limited in number, realism, and variability.⁷⁰ Although we thus could not investigate measurements with known ground truth, we believe that the WB-index is a good starting point to determine reproducibly the optimal number of clusters for [¹⁸F]FET-PET data.

Further points for improving clustering results

First, in the preprocessing step related to reduce the number of noisy voxels and extract the relevant voxels for further analysis, different thresholding methods, such as histogram shape-based methods can be implemented and compared. At the same time, different frame weighting and other reconstruction parameters schemes can be applied and compared with the one used in this study.

Second, Euclidean distance was used to assign each voxel-TAC to the nearest centroids. Different distance measures, as for example city-block, cosine and correlation distance measures^{72, 73}, can be implemented and compared with the one used in this study.

Third, in this study, clustering algorithm of TACs was performed on the temporal information without involving the spatial correlation between voxels. That is why for example the tumor voxels and some skin voxels belong to the same cluster (FIG.9 (upper row)). Thus, in future a segmentation method that utilizes both the spatial and temporal information in dynamic PET data may additionally aid in the evaluation of patients.^{34, 74}

Fourth, the optimal number of clusters for PET data can be determined based on non-least square fitting for generation of biological parameters. Accordingly, after PET data is being clustered, the resulting mean TACs can be used directly for estimating of kinetic parameters using a compartment model. A compartment is a volume in which the tracer is distributed uniformly. These models contain parameters which describe the transport rates of the tracer such as the flow in, flow out, and generally the uptake of the tracer inside specific cluster. These parameters can be estimated by fitting the model equations to TACs derived from PET images clustering. Then, the optimal number of clusters for PET data can be determined based on the biological

parameters derived from the compartments model such as all voxel-TACs belong to same cluster have similar biological parameters.

5. Conclusions

k -means clustering is a highly reproducible approach to define different functional structures on dynamic [^{18}F]FET-PET brain images compared to the manual drawing of ROIs for the investigated image reconstruction algorithm and the used scanner.

Among all investigated cluster validity indices, the WB-index could be used to reproducibly and automatically determine the optimal number of clusters for [^{18}F]FET-PET brain images.

In addition, the WB-index suggests a higher number of clusters compared to I , modified Dunn's, and Silhouette validity indices, thus allowing better differentiation of different brain tissues. Besides, the WB index has a higher reproducibility and a unique global minimum with a simple approach.

The methodology developed in this study is general and can also be performed on dynamic [^{18}F]FDG-PET images or any other types of tracer.

6. References

- 1 S. Goyal, T. Kataria, "Image guidance in radiation therapy: techniques and applications," *Radiol Res Pract* **2014**, 705604 (2014).
- 2 J. Bernier, E.J. Hall, A. Giaccia, "Radiation oncology: a century of achievements," *Nat Rev Cancer* **4**, 737-747 (2004).
- 3 M.M. Ter-Pogossian, "Basic principles of computed axial tomography," *Semin Nucl Med* **7**, 109-127 (1977).
- 4 R.A. Robb, "X-ray computed tomography: from basic principles to applications," *Annu Rev Biophys Bioeng* **11**, 177-201 (1982).
- 5 M. Pfeiler, K. Steiner, "[Computed tomography--basic principles and outlook for related procedures]," *Morphol Med* **3**, 117-124 (1983).
- 6 K.R. Maravilla, R.C. Murry, Jr., "Computed tomography: basic principles of operation," *Clin Neurosurg* **28**, 482-501 (1981).
- 7 C.E. Meacham, S.J. Morrison, "Tumour heterogeneity and cancer cell plasticity," *Nature* **501**, 328-337 (2013).
- 8 M. Hatt, F. Tixier, L. Pierce, P.E. Kinahan, C.C. Le Rest, D. Visvikis, "Characterization of PET/CT images using texture analysis: the past, the present... any future?," *Eur J Nucl Med Mol Imaging* 2016).
- 9 M. Aiello, C. Cavaliere, M. Salvatore, "Hybrid PET/MR Imaging and Brain Connectivity," *Front Neurosci* **10**, 64 (2016).
- 10 S. Basu, S. Hess, P.E. Nielsen Braad, B.B. Olsen, S. Inglev, P.F. Hoiland-Carlsen, "The Basic Principles of FDG-PET/CT Imaging," *PET Clin* **9**, 355-370, v (2014).
- 11 M.J. Martinez, S.I. Ziegler, T. Beyer, "PET and PET/CT: basic principles and instrumentation," *Recent Results Cancer Res* **170**, 1-23 (2008).
- 12 G. Omami, D. Tamimi, B.F. Branstetter, "Basic principles and applications of (18)F-FDG-PET/CT in oral and maxillofacial imaging: A pictorial essay," *Imaging Sci Dent* **44**, 325-332 (2014).
- 13 S. Watanuki, "[Attenuation correction methods in PET and PET/CT: basic principles and validity]," *Nihon Hoshasen Gijutsu Gakkai Zasshi* **62**, 797-803 (2006).
- 14 A.C. Clavo, R.L. Wahl, "Effects of hypoxia on the uptake of tritiated thymidine, L-leucine, L-methionine and FDG in cultured cancer cells," *J Nucl Med* **37**, 502-506 (1996).
- 15 C.C. Meltzer, J.P. Leal, H.S. Mayberg, H.N. Wagner, Jr., J.J. Frost, "Correction of PET data for partial volume effects in human cerebral cortex by MR imaging," *J Comput Assist Tomogr* **14**, 561-570 (1990).
- 16 S. Ziegler, "Basic physics behind new imaging techniques: Pet," *Radiother Oncol* **81**, S173-S173 (2006).
- 17 S.I. Ziegler, "Positron emission tomography: Principles, technology, and recent developments," *Nucl Phys A* **752**, 679c-687c (2005).
- 18 C.C. Meltzer, J.P. Leal, H.S. Mayberg, H.N. Wagner, J.J. Frost, "Correction of Pet Data for Partial Volume Effects in Human Cerebral-Cortex by Mr Imaging," *J Comput Assist Tomo* **14**, 561-570 (1990).
- 19 S.M. Bentzen, V. Gregoire, "Molecular Imaging-Based Dose Painting: A Novel Paradigm for Radiation Therapy Prescription," *Semin Radiat Oncol* **21**, 101-110 (2011).
- 20 N. Galldiks, G. Stoffels, C.P. Filss, M.D. Piroth, M. Sabel, M.I. Ruge, H. Herzog, N.J. Shah, G.R. Fink, H.H. Coenen, K.J. Langen, "Role of O-(2-(¹⁸F-fluoroethyl)-L-tyrosine PET for differentiation of local recurrent brain metastasis from radiation necrosis," *J Nucl Med* **53**, 1367-1374 (2012).
- 21 S.R. Cherry, J.A. Sorenson, M.E. Phelps, J.A. Sorenson, *Physics in nuclear medicine*, 3rd ed. (Saunders, Philadelphia, PA, 2003).

- 22 S.R. Cherry, J.A. Sorenson, M.E. Phelps, *Physics in nuclear medicine*, 4th ed. (Elsevier/Saunders, Philadelphia, 2012).
- 23 O. Warburg, "Note on the metabolism of tumours.," *Biochem Z* **228**, 257-258 (1930).
- 24 O. Warburg, F. Wind, E. Negelein, "The metabolism of tumors in the body.," *J Gen Physiol* **8**, 519-530 (1927).
- 25 K.J. Langen, K. Hamacher, M. Weckesser, F. Floeth, G. Stoffels, D. Bauer, H.H. Coenen, D. Pauleit, "O-(2-[¹⁸F]fluoroethyl)-L-tyrosine: uptake mechanisms and clinical applications," *Nucl Med Biol* **33**, 287-294 (2006).
- 26 N. Galldiks, G. Stoffels, C. Filss, M. Rapp, T. Blau, C. Tscherpel, G. Ceccon, V. Dunkl, M. Weinzierl, M. Stoffel, M. Sabel, G.R. Fink, N.J. Shah, K.J. Langen, "The use of dynamic O-(2-¹⁸F-fluoroethyl)-L-tyrosine PET in the diagnosis of patients with progressive and recurrent glioma," *Neuro Oncol* **17**, 1293-1300 (2015).
- 27 A. Habermeier, J. Graf, B.F. Sandhofer, J.P. Boissel, F. Roesch, E.I. Closs, "System L amino acid transporter LAT1 accumulates O-(2-fluoroethyl)-L-tyrosine (FET)," *Amino Acids* **47**, 335-344 (2015).
- 28 J.F. Cornelius, G. Stoffels, C. Filss, N. Galldiks, P. Sloty, M. Kamp, M. el Khatib, D. Hanggi, M. Sabel, J. Felsberg, H.J. Steiger, H.H. Coenen, N.J. Shah, K.J. Langen, "Uptake and tracer kinetics of O-(2-(¹⁸F)-fluoroethyl)-L-tyrosine in meningiomas: preliminary results," *Eur J Nucl Med Mol Imaging* **42**, 459-467 (2015).
- 29 J.A. Sorenson, M.E. Phelps, *Physics in nuclear medicine*. (Grune & Stratton, New York, 1980).
- 30 J.A. Sorenson, M.E. Phelps, *Physics in nuclear medicine*, 2nd ed. (Grune & Stratton, Orlando, 1987).
- 31 F. Molina-Duran, D. Dinter, F. Schoenahl, S.O. Schoenberg, G. Glatting, "Dependence of image quality on acquisition time for the PET/CT Biograph mCT," *Z Med Phys* **24**, 73-79 (2014).
- 32 M. Muzi, F. O'Sullivan, D.A. Mankoff, R.K. Doot, L.A. Pierce, B.F. Kurland, H.M. Linden, P.E. Kinahan, "Quantitative assessment of dynamic PET imaging data in cancer imaging," *Magn Reson Imaging* **30**, 1203-1215 (2012).
- 33 S. Mouysset, H. Zbib, S. Stute, J.M. Girault, J. Charara, J. Noailles, S. Chalon, I. Buvat, C. Tauber, "Segmentation of dynamic PET images with kinetic spectral clustering," *Phys Med Biol* **58**, 6931-6944 (2013).
- 34 J.X. Cheng-Liao, J.Y. Qi, "Segmentation of mouse dynamic PET images using a multiphase level set method," *Phys Med Biol* **55**, 6549-6569 (2010).
- 35 J. Lapuyade-Lahorgue, D. Visvikis, O. Pradier, C. Cheze Le Rest, M. Hatt, "SPEQTACLE: An automated generalized fuzzy C-means algorithm for tumor delineation in PET," *Med Phys* **42**, 5720-5734 (2015).
- 36 K.P. Wong, D.G. Feng, S.R. Meikle, M.J. Fulham, "Segmentation of dynamic PET images using cluster analysis," *IEEE T Nucl Sci* **49**, 200-207 (2002).
- 37 P. Velamuru, R. Renaut, H. Guo, K. Chen, "Robust Clustering of Positron Emission Tomography Data," in *Joint Conference of the Classification Society of North America and Interface Foundation of North America* (St. Louis, 2005).
- 38 J.C. Mazziotta, C.C. Pelizzari, G.T. Chen, F.L. Bookstein, D. Valentino, "Region of Interest Issues - the Relationship between Structure and Function in the Brain," *J Cerebr Blood F Met* **11**, A51-A56 (1991).
- 39 M. Hatt, C.C. Le Rest, N. Albarghach, O. Pradier, D. Visvikis, "PET functional volume delineation: a robustness and repeatability study," *Eur J Nucl Med Mol I* **38**, 663-672 (2011).
- 40 N.C. Krak, R. Boellaard, O.S. Hoekstra, J.W. Twisk, C.J. Hoekstra, A.A. Lammertsma, "Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial," *Eur J Nucl Med Mol Imaging* **32**, 294-301 (2005).

41 H. Zaidi, I. El Naqa, "PET-guided delineation of radiation therapy treatment volumes: a survey of
 image segmentation techniques," *Eur J Nucl Med Mol I* **37**, 2165-2187 (2010).

42 M. Halkidi, Y. Batistakis, M. Vazirgiannis, "On clustering validation techniques," *J Intell Inf Syst* **17**,
 107-145 (2001).

43 T. Kanungo, D.M. Mount, N.S. Netanyahu, C.D. Piatko, R. Silverman, A.Y. Wu, "An efficient k-
 means clustering algorithm: Analysis and implementation," *Ieee T Pattern Anal* **24**, 881-892
 (2002).

44 M. Liptrot, K.H. Adams, L. Martiny, L.H. Pinborg, M.N. Lonsdale, N.V. Olsen, S. Holm, C. Svarer,
 G.M. Knudsen, "Cluster analysis in kinetic modelling of the brain: a noninvasive alternative to
 arterial sampling," *NeuroImage* **21**, 483-493 (2004).

45 O. Arbelaitz, I. Gurrutxaga, J. Muguerza, J.M. Perez, I. Perona, "An extensive comparative study
 of cluster validity indices," *Pattern Recogn* **46**, 243-256 (2013).

46 Y.C. Liu, Z.M. Li, H. Xiong, X.D. Gao, J.J. Wu, S. Wu, "Understanding and Enhancement of Internal
 Clustering Validation Measures," *Ieee T Cybernetics* **43**, 982-994 (2013).

47 H. Akaike, "[Data analysis by statistical models]," *No To Hattatsu* **24**, 127-133 (1992).

48 T. Calinski, "A Dendrite Method for Cluster Analysis," *Biometrics* **24**, 207-& (1968).

49 L.J. Hubert, J.R. Levin, "General Statistical Framework for Assessing Categorical Clustering in
 Free-Recall," *Psychol Bull* **83**, 1072-1080 (1976).

50 Q.P. Zhao, P. Franti, "WB-index: A sum-of-squares based index for cluster validity," *Data Knowl*
Eng **92**, 77-89 (2014).

51 M. Soret, S.L. Bacharach, I. Buvat, "Partial-volume effect in PET tumor imaging," *J Nucl Med* **48**,
 932-945 (2007).

52 P.K. Saha, J.K. Udupa, "Optimum image thresholding via class uncertainty and region
 homogeneity," *Ieee T Pattern Anal* **23**, 689-706 (2001).

53 B. Foster, U. Bagci, A. Mansoor, Z.Y. Xu, D.J. Mollura, "A review on segmentation of positron
 emission tomography images," *Comput Biol Med* **50**, 76-96 (2014).

54 U. Maulik, S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and
 validity indices," *Ieee T Pattern Anal* **24**, 1650-1654 (2002).

55 N. Ilc, "Modified Dunn's cluster validity index based on graph theory," *Prz Elektrotechniczn* **88**,
 126-131 (2012).

56 P.J. Rousseeuw, "Silhouettes - a Graphical Aid to the Interpretation and Validation of Cluster-
 Analysis," *J Comput Appl Math* **20**, 53-65 (1987).

57 L. Jegatha Deborah, R. Baskaran, A. Kannan, "A Survey on Internal Validity Measure for Cluster
 Validation," *International Journal of Computer Science & Engineering Survey* **1**, 85-102 (2010).

58 C. Kratochwil, S.E. Combs, K. Leotta, A. Afshar-Oromieh, S. Rieken, J. Debus, U. Haberkorn, F.L.
 Giesel, "Intra-individual comparison of F-18-FET and F-18-DOPA in PET imaging of recurrent brain
 tumors," *Neuro-Oncology* **16**, 434-440 (2014).

59 G. Glatting, F.M. Mottaghy, J. Karitzky, A. Baune, F.T. Sommer, G.B. Landwehrmeyer, S.N. Reske,
 "Improving binding potential analysis in [¹¹C]raclopride PET studies using cluster analysis," *Med*
Phys **31**, 902-906 (2004).

60 M. Brun, C. Sima, J.P. Hua, J. Lowey, B. Carroll, E. Suh, E.R. Dougherty, "Model-based evaluation
 of clustering validation measures," *Pattern Recogn* **40**, 807-824 (2007).

61 A. Likas, N. Vlassis, J.J. Verbeek, "The global k-means clustering algorithm," *Pattern Recogn* **36**,
 451-461 (2003).

62 S. Stute, S. Vauclin, H. Necib, N. Grotus, P. Tylski, N.S. Rehfeld, S. Hapdey, I. Buvat, "Realistic and
 Efficient Modeling of Radiotracer Heterogeneity in Monte Carlo Simulations of PET Images With
 Tumors," *Ieee T Nucl Sci* **59**, 113-122 (2012).

- 63 D. Mahoney, S.C. Huang, A.R. Ricci, J.C. Mazziotta, R.E. Carson, E.J. Hoffman, M.E. Phelps, "A Realistic Computer-Simulated Brain Phantom for Evaluation of PET Characteristics," *IEEE Trans Med Imaging* **6**, 250-257 (1987).
- 64 S.K. Warfield, K.H. Zou, W.M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Trans Med Imaging* **23**, 903-921 (2004).
- 65 J.G. Brankov, N.P. Galatsanos, Y.Y. Yang, M.N. Wernick, "Segmentation of dynamic PET or fMRI images based on a similarity metric," *IEEE T Nucl Sci* **50**, 1410-1414 (2003).
- 66 R. Maroy, R.L. Boisgard, C. Comat, V. Frouin, P. Cathier, E. Duchesnay, F. Dolle, P.E. Nielsen, R. Trebossen, B. Tavitian, "Segmentation of rodent whole-body dynamic PET images: An unsupervised method based on voxel dynamics," *IEEE T Med Imaging* **27**, 342-354 (2008).
- 67 M.E. Kamasak, C.A. Bouman, E.D. Morris, K. Sauer, "Direct reconstruction of kinetic parameter images from dynamic PET data," *IEEE Trans Med Imaging* **24**, 636-650 (2005).
- 68 S. Jan, G. Santin, D. Strul, S. Staelens, K. Assié, D. Autret, S. Avner, R. Barbier, M. Bardiès, P.M. Bloomfield, D. Brasse, V. Breton, P. Bruyndonckx, I. Buvat, A.F. Chatzioannou, Y. Choi, Y.H. Chung, C. Comtat, D. Donnarieix, L. Ferrer, S.J. Glick, C.J. Groiselle, D. Guez, P.F. Honore, S. Kerhoas-Cavata, A.S. Kirov, V. Kohli, M. Koole, M. Krieguer, D.J.v.d. Laan, F. Lamare, G. Largeron, C. Lartzien, D. Lazaro, M.C. Maas, L. Maigne, F. Mayet, F. Melot, C. Merheb, E. Pennacchio, J. Perez, U. Pietrzyk, F.R. Rannou, M. Rey, D.R. Schaart, C.R. Schmidtlein, L. Simon, T.Y. Song, J.M. Vieira, D. Visvikis, R.V.d. Walle, E. Wieërs, C. Morel, "GATE: a simulation toolkit for PET and SPECT," *Physics in Medicine and Biology* **49**, 4543-4561 (2004).
- 69 S. Jan, D. Benoit, E. Becheva, T. Carlier, F. Cassol, P. Descourt, T. Frisson, L. Grevillot, L. Guigues, L. Maigne, C. Morel, Y. Perrot, N. Rehfeld, D. Sarrut, D.R. Schaart, S. Stute, U. Pietrzyk, D. Visvikis, N. Zahra, I. Buvat, "GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy," *Phys Med Biol* **56**, 881-901 (2011).
- 70 P. Papadimitroulas, G. Loudos, A. Le Maitre, M. Hatt, F. Tixier, N. Efthimiou, G.C. Nikiforidis, D. Visvikis, G.C. Kagadis, "Investigation of realistic PET simulations incorporating tumor patient's specificity using anthropomorphic models: Creation of an oncology database," *Medical Physics* **40**(2013).
- 71 S. Stute, T. Carlier, K. Cristina, C. Noblet, A. Martineau, B. Hutton, L. Barnden, I. Buvat, "Monte Carlo simulations of clinical PET and SPECT scans: impact of the input data on the simulated images," *Phys Med Biol* **56**, 6441-6457 (2011).
- 72 J.S. Chen, R.K.H. Ching, Y.S. Lin, "An extended study of the K-means algorithm for data clustering and its applications," *J Oper Res Soc* **55**, 976-987 (2004).
- 73 S.I. Abudalfa, M. Mikki, "K-means algorithm with a novel distance measure," *Turk J Electr Eng Co* **21**, 1665-1684 (2013).
- 74 W. Zhu, J. Ouyang, Y. Rakvongthai, N.J. Guehl, D.W. Wooten, G. El Fakhri, M.D. Normandin, Y. Fan, "A Bayesian spatial temporal mixtures approach to kinetic parametric images in dynamic positron emission tomography," *Med Phys* **43**, 1222 (2016).

7. Curriculum Vitae

Personal Information

Name : Bedor Abualhaj
Date of Birth : 21 July 1988
Place of Birth : Amman, Jordan
Marital Status : Married, one child

Education Background

2007-2008 High School, Ashamleh School, Amman, Jordan
(92 points out of 100)
2009-2010 Bachelor in Physics, Jordan University, Amman, Jordan
(3.95 points out of 4, rating: Excellent)
2012-2013 Master in Physics, Jordan University, Amman, Jordan
(4.00 points out of 4, rating: Excellent)
2014-now Doctoral student in Medical Radiation Physics/Radiation Protection, Medical
Faculty Heidelberg, Heidelberg University

Work Experience

2009-2013 Lecturer Assistant of Experimental Physics in Physics, Jordan University,
Amman, Jordan
2014-now Co-Supervision of one Master Theses in Biomedical Engineering and Medical
Physics Programs at Heidelberg University

Scholarships

2009-2010	Awarded a full scholarship from the University of Jordan for the Bachelor Degree Amman, Jordan
2012-2013	Awarded a full scholarship from the University of Jordan for the master Degree Amman, Jordan
2014-now	Awarded “Deutscher Akademischer Austausch-Dienst” (DAAD, German Academic Exchange Service, Germany, grant number: 91525125-57048249) scholarship for PhD study.

8. Publications and Abstracts

Articles

1. **Abualhaj B**, Weng G, Ong M, Attarwala AA, Molina-Duran F, Büsing K-A, Glatting G. Comparison of Five Cluster Validity Indices Performance in Brain [^{18}F]FET-PET Image Segmentation Using *k*-means. *Med Phys* 2017; 4: 209-220.
2. **Abu Alhaj B**, Hamad B. Electronic and magnetic properties of $\text{Co}_{2-x}\text{Fe}_x\text{CrAl}$ alloys: Ab initio calculations. *Phys. Status Solidi B* 2014; 251: 184–189.
3. **Abualhaj B**, Hamad B. The effect of defects on the electronic and magnetic properties of Co_2CrAl Heusler alloy. *J Phys Chem Solids* 2013; 74.
4. **Abu Alhaj B**, Hamad B, Khalifeh R, Shaltaf R. Ab-initio calculations of the electronic and magnetic structures of $\text{Co}_2\text{Cr}_{1-x}\text{Mn}_x\text{Si}$ alloys. *Journal of Magnetism and Magnetic Materials* 2013; 336: 37-43.
5. **Abualhaj B**, Hamad B. Ab-initio calculations of the electronic and magnetic structures of $\text{Co}_2\text{Cr}_{1-x}\text{Mn}_x\text{Al}$ alloys. *J Appl Phys* 2012; 112.

Abstracts

1. **B. Abualhaj**, G. Weng, A.A. Attarwala, M. Ong, F. Molina, K. Büsing, G. Glatting. [^{18}F]FET-PET brain image segmentation using *k*-means: Evaluation of four cluster validity indices. *Nuklearmedizin* 2017; 56.
2. **B. Abualhaj**, G. Weng, M. Ong, K. Büsing, G. Glatting. PET image clustering using *k*-means to identify tumor lesions for radiotherapy treatment planning. *DEGRO* 2016. 22. P19-5-jD.
3. **B. Abualhaj**, D. Hardiansyah, G. Weng, A.A. Attarwala, G. Glatting. Modelling the Arterial Input Function for Human Brain Studies. *Nuklearmedizin* 2016; 55: A50, V143.
4. **B. Abualhaj**, G. Weng, M. Ong, F. Molina-Duran, K. Buesing, G. Glatting. Investigation on Validity Indices for Brain [^{18}F]FET-PET Image Segmentation Using *k*-means Cluster Analysis. *DGMP* 2016 ;138, V55.

9. Acknowledgements

My work in this project gave me the chance to acquire more knowledge about the medical physics field. In addition, this work gave me the chance to know intelligent and diligent people who I would like to express for them my sincere gratitude. The prior thank is to the God for the good health and then to “Deutscher Akademischer Austausch-Dienst” for funding the project by a scholarship.

I would like to thank Professor Gerhard Glatting who gave me the opportunity to work under his guidance and support. His patience, motivation, and immense knowledge were the reasons to have nice results at the end. Thank you Professor for the wonderful environment you provided during the project. Absolutely, my words are not enough to express all my gratitude.

I also want to thank my husband Emad Mustafa for supporting me in all my life. He left his work in Jordan and accompanied me in my journey to Germany where we have our first baby (Jana).

I am using this opportunity to express my gratitude to everyone who supported me throughout my study. Special thanks to Mr. Ali Asgar Attarwala, Dr. Christian Maaß and Dr. Deni Hardiansyah for the stimulating discussions and for all the fun we had together. Also I thank Mr. Gouyang Weng, without him this thesis would not exist as it is.

Last but not the least; I am definitely grateful to my family: my parents, brothers and sisters for supporting me throughout my life in general. Thank you Mum because of your prayers I am always safe and I can always proceed.