Factor analysis-based approach for early uptake automatic quantification of breast cancer by $^{18}$F-FDG PET images sequence

Ines Ketata,a,b,*, Lamia Sallemi,a Frédéric Morain-Nicolier,b Mohamed Ben Slima,a Alexandre Cochet,d, Khalil Chtourou,a,c Su Ruan,c Ahmed Ben Hamida,a

a ATMS, Advanced Technologies for Medicine & Signals, Sfax University, 3038, Tunisia
b CReSTIC, Image Processing Group, REIMS University, 10026, France
c LITIS, Quantify, University of Rouen, 76183, France
d Le2i CNRS UMR 6306, Georges-François Leclerc Cancer Centre, Dijon, France
e CHU Sfax Hospital, Nuclear Medicine Service, Sfax University, Tunisia

ARTICLE INFO

Article history:
Received 12 March 2013
Received in revised form 12 June 2013
Accepted 29 July 2013

Keywords:
Breast cancer diagnostic
Dynamic PET images
Positron Emission Tomography (PET)
Segmentation
FAMIS approach
Quantification of tracer
First Pass Quantification ($K_{FPQ}$)

ABSTRACT

Factor Analysis of Medical Image Sequences (FAMIS) is recognized as one pioneer successfully used approach for analyzing especially dynamic images’ sequence for estimating kinetics and associated compartments having a physiological meaning.

Some studies tried to extend the exploring of this approach to analyze Positron Emission Tomography (PET) image modality for dynamic sequences. PET images with $^{18}$F-fluorodesoxyglucose ($^{18}$F-FDG) is the gold standard for in vivo, evaluation of tumor glucose metabolism and is widely used in clinical oncology.

In this paper, a novel approach is proposed to obtain an automated quantification method for early accumulation of $^{18}$F-FDG tracer in order to explore breast cancer, by applying FAMIS tool on dynamic first pass $^{18}$F-FDG PET dynamic sequences. This approach starts by an automated identification of a tumor Region of Interest (ROI) from PET dynamic images’ sequence. Then, a FAMIS approach is applied to separate two compartments: one compartment is associated to the vascular and a second one is associated to the purely tumor compartment. The latter allows the evaluation of the temporal evolution of the glucose tracer metabolism and therefore for pursuing cancer characterization.

A new empiric parameter $K_{FPQ}$ (First Pass Quantification), computed from the evolution of the $^{18}$F-FDG radiotracer accumulation using the first 11 min PET early images, is proposed. This parameter is found to be correlated to standardized uptake value maximal index (SUV$_{max}$) metabolism tumor.

The proposed framework is tested using image sequences’ database for 25 different pathology cases, which is considered as largely sufficient by the clinical team. Among clinicians’ experience, using a large dataset permits the possibility to obtain accurate information and precise early diagnosis. Pearson correlation coefficient is computed to evaluate as well as to analyze the relationship between the proposed empiric parameter $K_{FPQ}$ and glucose tracer metabolism SUV$_{max}$ for the overall pathology cases. $K_{FPQ}$ is successfully evaluated by the dynamic first-pass $^{18}$F-FDG PET image sequences for exploring early breast cancer diagnosis. Quantitative evaluations, as discussed and validated by clinicians, confirmed the efficiency of the modeling and the usefulness of the new empiric parameter $K_{FPQ}$ to predict tumor glucose metabolism for early uptake. This can be considered as a significant indication for quantification as well as evaluation of early relapse and disease progression during the therapy.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Today, medical technologies, especially the most recent and developed ones, are essential diagnostic tools, largely used in oncology explorations. The main objective behind such explorations is essentially cancer early detection before reaching adverse pathological serious stages. In fact, these advanced technologies are able to provide accurate anatomical information with static modalities such as Magnetic Resonance Imagery (MRI) images, and functional and/or metabolic information with dynamic modalities such as the functional Magnetic Resonance Imagery (fMRI) images, the Positron Emission Tomography (PET) image sequences and even the PET-Scan image sequences. Medical images functional processing is an active and recent domain that start to gain interest from clinicians. The aim of this type of images is to reinforce diagnosis and to offer

* Corresponding author at: ATMS, Advanced Technologies for Medicine & Signals, Sfax University, 3038, Tunisia. Tel.: +216 25 636 909.
E-mail address: ketata.ines22@gmail.com (I. Ketata).

1746-8094/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bspc.2013.07.008
important aids for clinical decision and/or for the conducting to pursue the therapy. In nuclear medicine, PET is considered as a very useful and effective tool, largely exploited in oncology explo-

rations. PET images’ sequences can provide significant information to be carefully studied for patient monitoring during treatment.

Several tools and approaches were proposed for the diverse functional modalities’ processing for several automatization tasks. In fact, among the clinical tasks, several manual ones are time consuming and not efficient for accurate diagnosis. Such manual tasks can be performed efficiently in automated way, and can be improved using powerful processing tools.

Several image analysis methods have been proposed in litera-

ture for medical diagnosis assistance in disease identification [1–5]. Among these methods, Factor Analysis of Medical Image Sequences (FAMIS) presents a particular interest for clinical studies to esti-

mate some physiological functions. Some publications discuss the clinical usefulness of this technique in the quantitative analysis of various types’ dynamic studies: gated cardiac [6,7] and renogra-

phy [8]. More recently, FAMIS was successfully used to analyze a dynamic sequence of scintigraphic images. It aims to estimate the variation of concentration of a tracer within different physiological functions [1–3,9–11].

In [12], FAMIS was used to analyze a dynamic sequence of medical images obtained via nuclear imaging modality, the PET images. It is based on a linear additive model, which assumes the under-

lying fundamental spatial distributions by factor images and the associated so-called fundamental functions (describing the signal temporal variations) by factors. It includes two steps: an orthogonal analysis and an oblique one. The goal of the first step is to determine a low-dimensional subspace, also called the ‘study subspace’ [13]. The second step is an oblique rotation of the basis vectors of the study subspace, in order to obtain non-orthogonal basis vectors, namely the factors, having a physical or physiological meaning.

PET images with $^{18}\text{F}-\text{FluorodeoxyGlucose}$ ($^{18}\text{F-FDG}$) is consid-

ered to be the gold standard for in vivo evaluation of tumor glucose metabolism. It is widely used in clinical oncology explorations. Different empirical studies have been proposed for breast cancer characterization. One example is the first-pass “Mullani” model. It is developed for extracting a measure of tumor perfusion from the first-pass delivery of $^{18}\text{F-FDG}$ to the tumor tissue manually seg-

mented. A first-pass 2 min scan of $^{18}\text{F-FDG}$ uptake (early uptake), after injection of the tracer, is used to compute blood flow (BF). Its derivation is based on the hypothesis that during the first tran-

sit of an activity bolus through the organ, there exists a duration where the tracer stay in tumor region, so that the venous concentra-

tion of the tracer is null. Another technique is proposed for tumor perfusion imaging based on immunohistochemical staining mea-

surement of angiogenesis with two antibodies to CD34 and CD105 used to evaluate tumor vascularization [14]. CD34 is a panendothelial marker that accurately reflects the degree of vascularization, whereas CD105 is a proliferation-related endothelial marker [15]. Tumor metabolism appears to be more associated with markers of proliferation, such as ki67. A standardized uptake value maximal index ($\text{SUV}_{\text{max}}$), reflecting tumor metabolism, is based on a static acquisition that is performed 90 min after injection (late uptake) for quantification of delayed $^{18}\text{F-FDG}$ tumor uptake [16]. It is con-

sidered as one of the most demanded parameter in breast cancer diagnosis.

In this paper, a novel approach for breast cancer exploration is proposed to perform an automated quantification of early accumu-

lation of $^{18}\text{F-FDG}$ parameter, by applying FAMIS tool on PET dynamic sequences. This approach starts by an automated identifi-

cation of a tumor Region of Interest (ROI) from PET dynamic images’ sequence. Then, FAMIS is applied to the tumor ROI to separate the two associated compartments: vascular and purely tumor. The latter allows the evaluation of the temporal evolution of the glucose

tracer metabolism and therefore for pursuing cancer characteriza-

tion.

In order to find a correlation between the obtained temporal tracer evolution from tumor compartment and $\text{SUV}_{\text{max}}$ parameter, we propose a novel empiric parameter, named $K_{\text{FAMIS}}$, characterizing the tumor glucose metabolism for early uptake. Pearson correlation coefficient is computed to evaluate and to analyze this correlation. Discussed and validated by clinician, the $K_{\text{FAMIS}}$ parameter can be considered as a significant indication for prediction of tumor glucose metabolism for early uptake, as well as an evaluation of early relapse and disease progression during the therapy.

The paper is structured as follows: in Section 2, FAMIS approach is firstly described as well as its performances and its limitations. Then, the methodology of PET images segmentation is presented. In Section 3, the tumor compartment parameterization is performed where an essential novel parameter $K_{\text{FAMIS}}$ and its mathematic formulation is proposed. In Section 4, experimental results, using semi-synthetic data as well as real-world ones obtained from 25 pathology cases’ data, are detailed. The validation of the proposed approach will be clinically discussed demonstrating clearly that the $K_{\text{FAMIS}}$ parameter yield a competitive marker for breast cancer explo-

rations. In Section 5, a conclusion on the efficiency of the proposed approach is conducted with work perspectives.

2. Advanced approach description

2.1. General methodology of the proposed approach

In this paper, a novel approach is proposed to obtain an auto-

mated quantification method for early accumulation of $^{18}\text{F-FDG}$ tracer in order to explore breast cancer using Factor Analysis of Medical Image Sequences (FAMIS) tool. FAMIS is a useful framework for automated breast cancer characterization with PET image modality, from the analysis of temporal dynamic first pass $^{18}\text{F-FDG}$ PET image sequences [17]. This approach starts by an automated identification of a tumor Region of Interest (ROI) from PET dynamic images’ sequence using an image segmentation method. Then, a FAMIS approach is applied on the obtained ROI to estimate the temporal evolution of the tracer concentration within two physi-

ological compartments: purely tumor and vascular. From the tumor compartment, one can then quantify the tracer glucose metabolism for early uptake and therefore for disease progression during the therapy.

2.2. Principle of FAMIS approach

FAMIS has been introduced in the eighties in nuclear medicine as a tool to summarize the relevant information underlying a dynamic or spectral image sequences by few images and associated curves, called factors. Indeed, FAMIS decomposes an images sequence, indexed by time, into few factor images and associated curves. The curve associated to any pixel in the sequence can be expressed as a linear combination of these factors. The weight associated with each factor corresponds to the pixel intensity in the associated factor image. Each factor image therefore represents the spatial distribution of the signal associated to a temporal curve [18,19].

Moreover, FAMIS uses some a priori information on the data model, such as the approximate number of physiological compart-

ments and the positivity of the factors on each compartment [11]. The mathematical model used to perform the decomposition of the dynamic image sequences is a linear superposition as described by the following equation [12,18]:

$$X = X(t) = \sum_{k} \mu_k(t) \cdot \eta_k(p) + \varepsilon(p),$$

(1)
where \( p \) denotes the pixel index and \( k \) is the index factor representing the physiological compartments number. This index factor should be relatively small and varying from 2 to 6. In fact, more this index is higher, less the analysis is robust [20]. The coefficients \( f_k \) coefficients are the fundamental kinetics, also called “factors”. The “factor image” \( f_k \) is the “factor image” associated to the \( k \)th physiological compartment. The parameter \( \varepsilon(p) \) parameter represents noise recorded for the \( p \)th pixel. As this linear model is always an approximation of real data, for instance because of noise, there is an error term in the model.

In this model, the kinetic \( X_p(t) \) for each pixel \( p \) is given by the input PET data, \( f_k \) and \( l_k \) are obtained as result of FAMIS analysis.

Curves recorded for each pixel are linear combination of a small number of fundamental kinetics having a physiological interpretation. For each pixel, the coefficient associated to each fundamental kinetic is the proportion of the signal in that pixel following the fundamental kinetic. The coefficients give the \( k \)th physiological compartment associated to the \( k \)th fundamental kinetic. The image of these coefficients (or weights) gives the compartment image with the kinetic described by the basic curve [11].

The input data is a sequence of \( M \) PET images of \( N \) pixels taken at different times referred to as \( t_1, t_2, \ldots, t_M \). \( X \) is the original data matrix of size \( N \times M \).

In this model, only kinetics are considered to be known. The physiological compartments and associated kinetics must be estimated. Moreover, noise recorded for each pixel had to be considered. To estimate all these unknown functions, the resolution method may use some a priori information related to noise properties.

2.3. FAMIS approach development

The resolution of the FAMIS model proceeds essentially in two major steps. The first one is an orthogonal analysis, used to estimate a low-dimensional subspace, also called the ‘study subspace’ [13,21,22], in which mainly of the relevant part of the pixels is represented, without the noise. The second step is an oblique analysis, used to determine physiological functions and associated compartments [11,21,22].

These two steps are preceded by a data preprocessing step performed on the input dynamic image sequences. In the following, we will give a brief description of each step.

2.3.1. Data preprocessing

This step aims to improve the quality of PET images and to increase the signal-to-noise ratio for each kinetic. We propose to apply a spatial mean of the slices that contains only the tumor region. The obtained preprocessed sequences of PET images will be further used for the next steps of FAMIS analysis.

2.3.2. Orthogonal analysis

A study subspace is determined in which all relevant information is supposed to be represented. It is important to well understand the notion of study subspace. The orthogonal analysis states that the observed signal is the sum of a fixed component (non noise-free component) and a random error. When the original data \( X_p(t) \) is a sequence of \( M \) images, it lies in a \( M \) dimensional space. However, there is often a vector of lower dimension \( X_p(t) - \varepsilon(p) \) belonging to a \( Q \) dimensional space \( S \) with \( Q \leq M \), in which all the relevant information underlying the original sequence is well represented. This is what we call the study subspace, represented by a plane, such that the projection of the original pixel onto this subspace gives the noise-free component.

The orthogonal analysis provides an estimation of the subspace \( S \). \( S \) is conventionally estimated by using the singular value decomposition (SVD) of data that diagonalizes the covariance matrix of \( X \), so that it minimizes the mean squared differences between the original data and the projection of the data on this space. The appropriate weights are chosen as a function of the statistical properties of the data. For instance, it has been shown that for PET data which are Poisson distributed, the appropriate orthogonal decomposition is the correspondence analysis [23].

In the study subspace \( X \) the relevant part of the data, can be written as a linear combination of the orthogonal basis vectors \( M \). Correspondence analysis (CA) yields an orthogonal spectral basis from which a \( Q \)-dimensional space \( S \), spanned by the \( Q \) eigenvectors associated with the \( Q \) largest eigenvalues of the covariance matrix decomposed by CA, is obtained [23].

The application of correspondence analysis is described through the steps below [23].

- Compute the following variables:

\[
x_i = \sum_{j=1}^{M} x_{ij}, \quad x_j = \sum_{i=1}^{N} x_{ij}, \quad x_\cdot = \sum_{j=1}^{M} x_{j}\cdot
\]

(2)

- Define the matrix \( y(N,M) \) issued from the optimal normalization of the data matrix \( X \) [23]:

\[
y_i = \frac{x_i}{x_\cdot}
\]

(3)

- Compute the following variables:

\[
o_q = \frac{x_i}{x_\cdot}, \quad \Gamma = \text{diag} \left( \frac{x_1}{x_\cdot}, \frac{x_2}{x_\cdot}, \ldots, \frac{x_M}{x_\cdot} \right).
\]

(4)

The mean vector \( y_{\text{moy}} \) is defined as

\[
y_{\text{moy}} = \left( \frac{2}{\sum_{i=1}^{N} o_i} \right) \sum_{i=1}^{N} o_i y_i = \left( \frac{x_1}{x_\cdot}, \frac{x_2}{x_\cdot}, \ldots, \frac{x_M}{x_\cdot} \right).
\]

(5)

The covariance matrix is defined as

\[
C = (y - 1_N \cdot y_{\text{moy}})' \cdot W \cdot (y - 1_N \cdot y_{\text{moy}}) \cdot M,
\]

(6)

with

\[
M = \Gamma^{-1} \quad \text{and} \quad W = \text{diag} \left( \frac{x_1}{x_\cdot}, \frac{x_2}{x_\cdot}, \ldots, \frac{x_M}{x_\cdot} \right).
\]

(7)

- Compute the \( Q \) eigenvectors \( u_q \) associated with the \( Q \) largest eigenvalues \( \lambda_q \) of the matrix \( C \) [23]. The \( V_Q \) is computed using the following equation:

\[
V_Q = (y - 1_N y_{\text{moy}}) \cdot M \cdot U_Q \cdot A_Q^{-1},
\]

(8)

with

\[
A_Q = \text{diag} \left( \sqrt{\lambda_1}, \ldots, \sqrt{\lambda_Q} \right).
\]

(9)

Each free-noise kinetic may be expressed as

\[
y_i(t) = y_{\text{moy}}(t) + \sum_{q=1}^{Q} \sqrt{\lambda_q} \cdot y_q(t) \cdot u_q(t).
\]

(10)

The images from CA contain specific negative values. The eigenvectors have also positive and negative values. The linear combination of these weights and curves is difficult to physiologically interpret. An oblique analysis allows introducing positivity constraints.
2.3.3. Oblique analysis  

The second step of FAMIS consists in finding $K$ non-orthogonal basis vectors, namely the factors, having a physical or physiological meaning in the subspace $S$ resulting from the orthogonal analysis. To locate these factors, physiological or physical priors related to the factors and the associated images are used as constraints in an iterative procedure that minimizes the discrepancy between the model and the data based on constraints of the positivity of factors and factor images [24].

Once the factors $f_i$ are identified, the negative $f_i$ are divided by two and normalized. This process is iteratively repeated and is to be stopped when the number of negative factors $f_i$ is considered to be low enough [3]. Oblique analysis is described through the steps below:

- Rewrite Eq. (10) as
  \[ y_i(t) = \sum_k q_k(i) f_k(t) \quad \text{with} \quad q_k(i) > 0 \quad \text{and} \quad f_k(t) > 0. \]  

- Orthogonal analysis gives a noise-free data matrix described as [25]:
  \[ Y = 1_N \cdot Y_{moy} + V_Q \cdot A_Q \cdot U_Q. \]  

- Oblique analysis aims to represent $Y$ as a combination of factors (matrix $F$) and factorial images (matrix $A$) with respect to the constraints ($A > 0$ and $F > 0$):
  \[ Y = A \cdot F \quad \text{with} \quad A > 0 \quad \text{and} \quad F > 0. \]

- The algorithm focuses on finding the transformation matrix $T$ that resolves [25]
  \[ F = T \cdot U \quad \text{with} \quad A > 0 \quad \text{and} \quad F > 0. \]  

The matrix $T$ is defined as

\[
     T = \begin{bmatrix}
        1 \\
        \vdots \\
        T_Q \\
        1
     \end{bmatrix}. 
\]  

First, constraints are applied to images. Then, the associated curves are deduced by projection. This procedure is repeated until the convergence of the algorithm to a stable solution [25].

The steps of the iterative algorithm are as follows:

1. Initialization of the matrix $T$.
2. Computation of the factorial images $A$
   \[
   A = V \cdot \Delta \cdot T^{-1}. 
   \]
3. Modify the values $q_k(i)$ using the constraints of positivity and normalization.
4. Computation of $T$ using the relation
   \[
   T^{-1} = A^{-1} (V' DV)^{-1} V'DA. 
   \]
5. Determination of $F$ using Eq. (14).
6. Modify the values $f_k(t)$ using positivity and normalization constraints.
7. Computation of $T$ using the relation
   \[
   T = FMU(U^T MU^T)^{-1}. 
   \]
8. Test of the stopping condition and return to the first step if it is not verified.

2.4. FAMIS applied to region of interest (ROI)

Previous study [12,24] has shown that the factor analysis applied to a sequence of dynamic PET images may not be discriminative to assess the dynamic time evolution of the tracer, so to specify the components of a tumor. To better understand the importance of the factor analysis, we try to analyze its specificities compared to the well-known areas of interest. The Region of Interest (ROI) selection has two advantages: the first is to allow a good extraction of physiological compartments. The second is an automated quantification for early accumulation parameter of $^{18}$F-FDG in breast cancer based only on the tumor region of first pass $^{18}$F-FDG PET dynamic sequences to cancer characterization. It is be described next.

Several methods have been proposed in literature [26,27] to determine an ROI. Among them there are the standard active contours, watersheds, level-set, etc. The latter method is widely used in image segmentation due to its capacity to integrate the detection procedure and the contour chain in the process of energy minimization. This method is also characterized by its lower computational complexity. For these reasons, the level-set method had been selected for the estimation of the tumor ROI on each image of the sequence [27]. The principle of level-set is to propagate a smooth continuous closed curve and generate a family of curves obtained by the movement of the curve in its normal speed of propagation, which depends on its curvature [26]. This segmentation is performed for each image sequence to construct finally a new sequence that contains only ROI areas.

Each ROI of an image corresponds to the tracer activity in the tumor. It corresponds to the summation of the activities of a number of anatomical and functional structures. Hence, the time-curves associated to the ROI time is a mixture of fundamental curves, each having a physiological interpretation.

Linear superposition model used by FAMIS is justified from a physiological point of view. This analysis can separate sources if a priori information on the number of physiological compartments is available. The determination of this number is useful. It depends on the studied body and it is based on its physiology.

In our medical cases study, the clinical team provides the number of physiological compartments. They fix this number to two. In fact, the determination of the number of compartments is based on the physiology of the repartition of the tracer used (FDG) in biologic tissue. Given the fact that this tracer is injected intravenously, its first appearance in the ROI reflects vascular distribution (vascular compartment). Then the tracer will continue its flow to the tumor cells, and its degree of retention reflects the tumor metabolism (tumor compartment). FAMIS approach applied to each ROI area extracts separately the vascular, as well as the purely tumor compartments and evaluates their activities and variations within time. The FAMIS approach with different number $k$ becomes less robust, i.e., we obtain the curves that have no physiological interpretation.

From FAMIS results, the input image sequence in the tumor ROI is decomposed as:

\[
   X_p(t) = \sum_{k=1}^{2} f_k(t) \cdot I_k(p). 
\]

This decomposition is reduced to two factors:

\[
   X_p(t) = A_p(t) + T_p(t), 
\]

with

\[
   A_p(t) = f_0(t) \cdot I_0(p). 
\]
being the arterial term (vascular term) and $T_p(t)$ being the tumor term:

$$T_p(t) = f_1(t) \cdot I_1(p).$$

3. Tumor compartment parameterization

3.1. Proposed model parameter: early uptake for quantification by $^{18}$F-FDG PET images sequence

Among clinical results, malignant tumors are generally associated with a noticeable increase of glucose metabolism. Such property has to be explored clinically using efficient processing tools applied to PET image modality. Thus, a static acquisition is performed 90 min after $^{18}$F-FDG tracer injection for the quantification of delayed $^{18}$F-FDG tumor uptake ($\text{SUV}_{\text{max}}$), reflecting tumor metabolism [16]. In locally advanced breast cancer, increased glucose metabolism predicts a poor response to neoadjuvant chemotherapy. The result is that the $\text{SUV}_{\text{max}}$ linked to a tumor is considered as one of the most demanded parameter that can be used to characterize breast cancer response to early relapse and disease progression.

The count rate measurement of a tracer in tumor with the use of dynamic models allows the glucose metabolism quantification. The use of a model can then track the tracer temporal evolution and thus the tumor metabolism. An automated approach is proposed using factor analysis tool to extract a new parameter computed from the two obtained compartments in the tumor ROI as evaluated by first pass $^{18}$F-FDG in PET early images (the first 11 min). The aim is to find a possible correlation between the parameter resulting from tumor compartment and the physiological parameter $\text{SUV}_{\text{max}}$ to reflect tumor glucose metabolism of a tracer for early uptake.

The number of compartments is a priori information that could provide by the clinicians. If one can suppose the existence of two compartments in the tumor ROI (vascular and tumor), then the parameter $k$ should be fixed to 2 (see Eq. (1)). Factor Analysis will provide the tracer temporal-evolution of the vascular compartment $A(t)$ as well as of the tumor compartment $T(t)$. In order to make further analysis of this tracer temporal-evolution, one can suppose the synthetic curves displayed in Fig. 1. One can assume that the factor that accumulates during the first time of the acquisition is linked to the vascular passage of the tracer. Whereas, the factor stabilized at the end of the acquisition is rather related to the tissue factor.

This model postulates that during the first transit of the tracer through tissue, the tracer is retained for an amount of time in which the tracer concentration in the venous is very small. This time delay ($t_s$) [28] is longer with highly extracted tracers such as $^{18}$F-FDG. At $t_s$, almost all tracer quantity is retained in the ROI, and so the venous outlet is close to zero. When $t$ is less than $t_s$, the counting rate reaches a maximum when vascular input to the ROI ceases; both this input and the venous drainage are then zero [29]. This calculation is valid at $t_m$ (peak-count time), the time when the amount of tracer in the tumor is the largest, i.e. at the end of the first passage [30].

In order to evaluate the tracer glucose metabolism for early uptake, we propose a new empiric parameter evaluated through the dynamic first-pass $^{18}$F-FDG PET in the setting of newly diagnosed breast cancer. The general first-pass new empiric parameter $K_{FPQ}$ can be adapted to the above-mentioned peak-count time $t_m$. This parameter is given by the following equation:

$$K_{FPQ} = \sum_{i=1}^{t_m-t_0} I_1(p) \cdot \frac{d}{dt}(f_1(t)) - \sum_{i=t_0+1}^{t_m} I_1(p) \cdot \frac{d}{dt}(f_1(t)).$$

where $p$ is a pixel in the image, $t_m$ (peak-count time) is the time at the end of the first passage, $I_1(p)$ and $f_1(t)$ are respectively the factor image and factors associated to the tumor compartment as result of FAMIS analysis (see Eq. (22)). According to Eq. (23), we remark that only the tracer activity occurring in the $K_{FPQ}$ parameter. This latter is not influenced by the body weight.

This model characterizes the difference between the slope (derivative of the function $f_1(t)$) of the time evolution of tracer when almost all tracer quantity is retained in the tumor compartment, (i.e. when the counting rate of the tracer will reach a maximum) and the slope of this tracer when it leaves the tumor.

3.2. Algorithm: code recapitulation

We recapitulate the processing operations of the proposed approach. It starts by an automated identification of a tumor Region of Interest (ROI) from PET dynamic images sequence using an image segmentation method. Then, a FAMIS approach is applied on the obtained ROI to estimate the temporal evolution of the tracer concentration within two physiological compartments: purely tumor and vascular. The resolution of the FAMIS model proceeds essentially in two major steps. The first one is an orthogonal analysis used to estimate the k factors number. The second step is an oblique analysis, used to determine physiological functions and associated compartments. From the tumor compartment, one can then quantify the tracer glucose metabolism for early uptake and therefore for disease progression during the therapy.

4. Experimental results and discussion

4.1. Synthetic data validation

Before considering real pathological cases, we discussed carefully with clinical team some possible synthetic data that can involve and describe real cases with several logical variations. For doing so, we had specifically discussed the possible tumor shape that can really occur regarding the speckled form generally depicted in malignant pathological cases and the curved form generally depicted in benign tumors’ pathological cases. Another criterion was also carefully discussed which was the tumors’ size which can be sometimes even small size. In our developed approach, and among the clinical team, it was sufficient to
distinguish carefully one small size region, and this can be considered as a very important result.

4.1.1. Synthetic data validation regarding the shape and the size

Fig. 2 shows an example of sequence $X_p(t)$ of synthetic images that contains speckle noise. This sequence is obtained as follows:

$$X_p(t) = f_1(t) \cdot I_1(p) + f_2(t) \cdot I_2(p) + \epsilon(p).$$  (24)

The square shape resembles the possible tumor shape that can really occur regarding the speckled form generally depicted in malignant pathological cases and the curved form such as the disk sharp generally depicted in benign tumors’ pathological cases. Another criterion was also carefully discussed which was sufficient to distinguish one small size region that corresponds to tumors' size.

The results of factor analysis show that there are three compartments. Two are displayed in Fig. 3: the disk and the square, the noise is not represented. Each compartment is associated with a physiological function. This confirms that we have a good correspondence in terms of spatial and temporal between the original data and the result obtained using the factor analysis approach.

It is known that the most limitations of PET imaging reside in its low spatial resolution and SNR. In order to take into account the Partial Volume Effect (PVE) [31] in the simulation study, a low-pass Gaussian filter is used to generate the PVE. Moreover, in order to vary the degree of the PVE, we applied the Gaussian filter on original image sequence $M$ times ($M$ varying from 1 to 3). In fact, the Gaussian filter relates to spatial resolution effect on visual appearance of boundaries between different regions of the image mainly when number iteration of filtering increases.

Figs. 4, 6 and 8 show the synthetic images simulating the PVE using Gaussian filter of size $3 \times 3$ with $M = 1, 2$ and 3 respectively. We observe that the imaged structures profile is blurred and spread as well as the object boundaries are made difficult to discern.

The results of factor analysis show that there are two compartments, as displayed in Figs. 5, 7 and 9. Each compartment is associated with a physiological function. This confirms that a good correspondence is obtained between the original data and the results obtained using the FAMIS approach with the PVE in terms of spatial and temporal (Figs. 6–9).

4.1.2. Synthetic data validation regarding the overlap and the noise

For quantitative evaluations, we use a mean absolute error and Cohen’s Kappa. The index Kappa ($IK$) measures the ratio overlap between the segmented image and ground truth as follows:

$$IK = \frac{2VP}{2VP + FN + FP}.$$  (25)

with $VP$, $FN$, and $FP$ are respectively the True Positives, False Negatives and False Positives number.

The reliability of source overlap in absence of noise is evaluated by keeping the disk fixed and translating the square from left to right. In order to evaluate the robustness of the proposed approach, a speckle noise is added to the data with three different variance values $\text{variance} = 0.5, 1$ and $1.5$. We assume that the range of these variance values corresponds to a typical variance of real PET images sequence. In fact, for the real-world PET images sequence used in this study, we evaluate the variance in the tumor ROIs whose the intensity is normalized between 0 and 1. We found a maximum value of order 0.1393 in these tumor ROIs of each image of the PET sequence for the 20 patients. Moreover, for the chosen variance values of speckle noise, we evaluate the variance of the noisy region of the synthetic images and found a maximum values of 0.1286, 0.1450 and 0.1693 for $\text{variance} = 0.5, 1$ and $1.5$ respectively.

The maximum overlap corresponds to position 6 (see Fig. 10). One can see that the best performances are reached when the overlap decreases, i.e. the index kappa is high and oppositely with the mean absolute error. This confirmed that our approach is reliable on the context of sources separation.

According to Fig. 10, when noise variance increases, source separation performance is similar. This confirms the denoising effect after the factor analysis application.

From the mean absolute error curve (see Fig. 10), one can observe a maximum mean error of 0.11. This corresponds to 10% of amplitude data that appears reasonable. The mean error did not give an idea on the spatial distribution of the information with respect to the ground truth that is why the kappa index is computed. This index is equal to 0.74 at the maximum overlap (see Fig. 10). This confirms that we have a good correspondence between the original shape and the shape obtained using the factor analysis approach ($kappa$ index $\geq 0.7$).

These interpretations confirm the reliability of our approach since it has proved better robustness to noise variance (noise effect is the same) and robustness to high overlap. These criteria are adapted to PET images.

4.2. Real-world data validation: pathology cases

From March 2010 to August 2011, 20 patients (mean age, 46 years; range, 27–72 years) with invasive breast cancer were prospectively recruited for $^{18}$F-FDG PET before any treatment. The inclusion criterion was newly diagnosed breast cancer, with primary tumor diameter larger than 2 cm. The exclusion criteria were pregnant or nursing patients, patients unwilling or unable to undergo serial imaging studies, and patients with inflammatory tumors. The clinical stage, determined according to the TNM classification of malignant tumors, was assessed by clinical examination, mammography, ultrasonography, pulmonary chest radiography, bone scintigraphy, and $^{18}$F-FDG PET images. These patients were included in clinical trials with the main objective to determine the pathological response at surgery; an ancillary study with $^{18}$F-FDG PET was planned and included within the informed consent.

4.2.1. PET images sequence example

For experimental tests, a sequence of FDG-PET images, provided by LECLERC medical center (FRANCE), is used. This sequence represents a breast tumor tissue exam. It contains a set of 540 DICOM files including 11 data volumes, each of 45 slices.

For the validation of our approach, the clinical team has selected some slices out of 45 that contain the tumor region. Some example slices are displayed in Fig. 11.

4.2.2. Results of FAMIS application on tumor ROI

In order to determine the tumor ROI, a shape of a circular object was used in the level-set segmentation method. Fig. 12 shows the converged result of the evolution of the level-set function. One can observe that, through the segmentation algorithm iteration (100 iterations are enough to achieve a good convergence), the evolution level-set function fits efficiently the tumor ROI.

Fig. 13 presents the results of factor analysis applied to tumor ROI. One can see that two different compartments are obtained:

- **Vascular compartment $A(t)$**: the associated physiological curve is decreasing over time, which means that the radiotracer accumulates on first time in the aorta compartment.
- **The tumor compartment $T(t)$**: its physiological curve increases over time before being stabilized. This is mainly due to the accumulation of the radiotracer in this compartment at the end of the acquisition.
Fig. 2. Construction of synthetic images sequence. Left: $I_1, f_1$ Image & Factor of disk compartment, $I_2, f_2$ Image & Factor of square compartment. Right: An example of sequence that contains 20 grayscale synthetic images corrupted by speckle noise. Each image is a mixture of different compartments.

Fig. 3. Results of factor analysis: (a and c) Image & Factor of disk compartment and (b and d) Image & Factor of square compartment.

Fig. 4. Partial volume effect simulation using Gaussian filter of size $3 \times 3$. $I_1$ is the disk compartment and $I_2$ is the square compartment.
4.3. Evaluation of 18F-FDG proposed model parameter for evaluation of SUV using one first pass model

The proposed framework has been tested on 25 patients. Only 20 pathology cases are retained to have visible breast cancer. Using a large dataset gives us the possibility to obtain accurate information and precise diagnosis in the setting of newly diagnosed breast cancer.

In this section, we evaluate the new proposed empiric parameter \( K_{p0q} \) based on the obtained tumor compartment curve, will be evaluated.

Fig. 14 shows typical curves of the vascular time-activity measured over the aorta (drawn with a dotted line) and the tumor uptake of 18F-FDG. It shows that total first-pass activity has been delivered to the tumor compartment at peak-counts time \( t_m \). This relationship between peak counts and vascular input function is used to verify that, at \( t_m \), venous concentration can be approximated to zero since \( A_p(t_m) \) is close to zero. These curves are used to measure the peak-count time \( t_m \). The variable \( t_m \) is fixed empirically by superimposing together the previous two curves (see Fig. 14). The time \( t_m \) is the one when the venous outlet is close to zero and the counting rate of the tracer reaches a maximum value in the tumor compartment. For the studied cases, this time was equal to 10 min.

Table 1 gives the values of the proposed empiric parameter \( K_{p0q} \) computed according to Eq. (23), as well as of the parameters SUV\(_{\text{max}}\), Ki67, CD34, CD105 and BF for the different studied patients.

For the standardized uptake value maximal index (SUV\(_{\text{max}}\)) parameter, it is computed according to the following equation:

\[
\text{SUV}_{\text{max}} = \frac{C(t)}{\text{injected dose}/\text{body weight}},
\]

where \( C(t) \) is the blood concentration of 18F-FDG and others parameters given by the clinical team before neoadjuvant chemotherapy.

In order to find a possible correlation between the empiric parameter \( K_{p0q} \) before chemotherapy and the parameters SUV\(_{\text{max}}\), Ki67, CD34, CD105 and BF, Pearson correlation coefficient is computed to evaluate and analyze the correlation. All tests were 2-sided, and a \( P \) value of less than 0.05 was considered significant. Analyses were performed using the SPSS software package (version 13.0; SPSS Inc.).

For overall patient population, it was found that there was a strong significant positive correlation between the parameters.
Fig. 7. Results of FAMIS applied to the synthetic images sequence with PVE using twice times Gaussian filter of size $3 \times 3$: (a and c) Image & Factor of square compartment and (b and d) Image & Factor of disk compartment.

Fig. 8. Partial volume effect simulation using three times Gaussian filter of size $3 \times 3$: $I_1$ is the disk compartment and $I_2$ is the square compartment.

Table 1

<table>
<thead>
<tr>
<th>Patients</th>
<th>$K_{FEO}$</th>
<th>$SUV_{max}$</th>
<th>Ki67</th>
<th>CD34</th>
<th>CD105</th>
<th>BF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>-0.31</td>
<td>5.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 2</td>
<td>281.14</td>
<td>5.42</td>
<td>80</td>
<td></td>
<td>320</td>
<td>214</td>
</tr>
<tr>
<td>Patient 3</td>
<td>205.30</td>
<td>11.40</td>
<td>80</td>
<td>194</td>
<td>137.04</td>
<td>0.32</td>
</tr>
<tr>
<td>Patient 4</td>
<td>417.75</td>
<td>6.48</td>
<td>40</td>
<td></td>
<td>50</td>
<td>126</td>
</tr>
<tr>
<td>Patient 5</td>
<td>352.67</td>
<td>2.28</td>
<td>20</td>
<td>88</td>
<td>352.67</td>
<td>0.15</td>
</tr>
<tr>
<td>Patient 6</td>
<td>29.88</td>
<td>5.56</td>
<td>20</td>
<td>284</td>
<td>52</td>
<td>0.57</td>
</tr>
<tr>
<td>Patient 7</td>
<td>85.23</td>
<td>7.03</td>
<td>20</td>
<td>214</td>
<td>94</td>
<td>0.38</td>
</tr>
<tr>
<td>Patient 8</td>
<td>117.92</td>
<td>5.17</td>
<td>50</td>
<td>74</td>
<td>46</td>
<td>0.3</td>
</tr>
<tr>
<td>Patient 9</td>
<td>15.35</td>
<td>5.50</td>
<td>60</td>
<td>174</td>
<td>75</td>
<td>0.23</td>
</tr>
<tr>
<td>Patient 10</td>
<td>200.87</td>
<td>5.10</td>
<td>70</td>
<td>116</td>
<td>58</td>
<td>0.24</td>
</tr>
<tr>
<td>Patient 11</td>
<td>434.38</td>
<td>10.40</td>
<td>80</td>
<td>184</td>
<td>56</td>
<td>0.25</td>
</tr>
<tr>
<td>Patient 12</td>
<td>513.78</td>
<td>9.28</td>
<td>60</td>
<td>130</td>
<td>513.78</td>
<td>0.46</td>
</tr>
<tr>
<td>Patient 13</td>
<td>120.52</td>
<td>6.79</td>
<td>30</td>
<td>395</td>
<td>92</td>
<td>0.36</td>
</tr>
<tr>
<td>Patient 14</td>
<td>612.56</td>
<td>9.18</td>
<td>60</td>
<td>26</td>
<td>44</td>
<td>0.21</td>
</tr>
<tr>
<td>Patient 15</td>
<td>125.79</td>
<td>6.52</td>
<td>60</td>
<td>142</td>
<td>84</td>
<td>0.33</td>
</tr>
<tr>
<td>Patient 16</td>
<td>1330.40</td>
<td>14.22</td>
<td>80</td>
<td>162</td>
<td>48</td>
<td>0.47</td>
</tr>
<tr>
<td>Patient 17</td>
<td>898.74</td>
<td>13.04</td>
<td>35</td>
<td>66</td>
<td>60</td>
<td>0.17</td>
</tr>
<tr>
<td>Patient 18</td>
<td>52.88</td>
<td>8.39</td>
<td>90</td>
<td>96</td>
<td>118</td>
<td>0.38</td>
</tr>
<tr>
<td>Patient 19</td>
<td>490.25</td>
<td>14.53</td>
<td>100</td>
<td>138</td>
<td>92</td>
<td>0.49</td>
</tr>
</tbody>
</table>
Fig. 9. Results of FAMIS applied to the synthetic images sequence with PVE using three times Gaussian filter of size $3 \times 3$: (a and c) Image & Factor of square compartment and (b and d) Image & Factor of disk compartment.

Fig. 10. Validation metric obtained through the variation of noise level: noise with $\text{mean} = 0$ and $\text{variance} = 0.5$ (red line), noise with $\text{mean} = 0$ and $\text{variance} = 1$ (blue line), noise with $\text{mean} = 0$ and $\text{variance} = 1.5$ (dotted blue line). Left: Mean absolute error. Right: Kappa index. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Fig. 11. Several 18-FDG PET breast cancer images for the same patient. An example of a few slices with a spatial mean that contains only the tumor region at different times.
Factor to obtained peak concentration aorta

Fig. 12. Results of the evolution of level-set segmentation function. Left: Initialization of the curve around the tumor region. Right: Results of the tumor segmentation obtained by the level-set: the tumor region is surrounded with a red line. The evolution level-set function fits efficiently the tumor ROI. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

(a) (b)

(c) (d)

Fig. 13. FAMIS result to a ROI of a PET sequence (images are restricted to tumor ROI to improve visibility). (a) and (c) Image & Factor of tumor compartment, (b) and (d) Image & Factor of vascular compartment.

Fig. 14. Temporal relationship between vascular concentrations as measured over aorta and tumor uptake of $^{18}$F-FDG. It shows that total first-pass activity has been delivered to tumor compartment at peak-counts time $t_m$. This relationship between peak counts and vascular input function is used to verify that, at $t_m$, venous concentration can be approximated to zero since $A_p(t_m)$ is close to zero.

$SUV_{\text{max}}$ and $K_{\text{FPQ}}$ ($r = 0.706$ with test value $P = 0.002$). In contrast, there was no significant correlation between $K_{\text{FPQ}}$ and ki67 score, neither for CD34 score nor for CD105 score, nor for BF score ($r = 0.14$, $r = 0.38$, $r = 0.01$ and $r = 0.17$, respectively, $P$ was not statistically significant for either). Fig. 15 shows the correlation between the parameters $SUV_{\text{max}}$ and $K_{\text{FPQ}}$, where one can observe the obtained linear correlation. We did not see any interest to present the correlation coefficients for the other parameters since there was no significant correlation obtained.

The new empiric parameter $K_{\text{FPQ}}$ is analogous to the metabolic rate of $^{18}$F-FDG uptake tumor (late uptake) presented in our work, which predicts tumor glucose metabolism for early uptake for the values before chemotherapy. On the other side, the parameter $SUV_{\text{max}}$, which is related to tumor metabolism, is determined on delayed PET images performed 90 min after injection. The manner to determine this parameter and PET exam itself is invasive. Thus, it would be worthwhile to identify a marker of metabolism determined on early images, immediately after injection. $K_{\text{FPQ}}$ can be considered as a significant indication for quantification as well as evaluation of early relapse and disease progression during the therapy. This allows to save time and to convenience to the patient.
5. Conclusion

An early uptake advanced approach for breast cancer exploration was proposed in this paper. Our main purpose in this research was to perform an automatic quantification method for early accumulation of $^{18}$F-FDG parameter, which could be an essential clinical criterion permitting to diagnose and to explore efficiently breast tumors. This was hence performed by applying FAMIS tool on dynamic first pass $^{18}$F-FDG PET dynamic sequences. This approach starts by an automated identification of the tumor Region of Interest (ROI) from PET dynamic images’ sequence. Then, a FAMIS approach was applied to separate the two associated compartments: one compartment was associated to the purely tumor region and a second one was associated to the vascular compartment. The tumor compartment could be useful for evaluating the temporal evolution of the glucose tracer metabolism and therefore for pursuing cancer characterization.

A new empirical parameter $K_{FPQ}$ (First Pass Quantification) was proposed in our proposed approach, which was found to be linked to standardized uptake value maximal index ($SUV_{max}$) metabolism tumor, and computed from the evolution of the $^{18}$F-FDG radiotracer accumulation. This was evaluated using the first 11 min PET early images.

Before considering real pathological cases, we discussed carefully with clinical team some possible synthetic data that could involve and describe real cases with several logical variations. For doing so, we had specifically discussed the possible tumor shape that could really occur regarding the speckled form generally depicted in malignant pathological cases and the curved form generally depicted in benign tumors’ pathological cases. Another criterion was also carefully discussed which was the tumors’ size which could be sometimes even small size. In our developed approach, and among the clinical team, it was sufficient to distinguish carefully one region up to one small size, and this could be considered as a very important result.

Image sequences’ database for 20 different pathology cases was successfully tested in our proposed framework. Such data discussed and furnished by the clinical team could be considered as largely sufficient. Among clinicians’ experience, using a large dataset permits the possibility to obtain accurate information and precise early diagnosis. Pearson correlation coefficient was computed to evaluate as well as to analyze the relationship between the proposed empiric parameter $K_{FPQ}$ and glucose tracer metabolism $SUV_{max}$ for the overall pathology cases. $K_{FPQ}$ was successfully evaluated by the dynamic first-pass $^{18}$F-FDG PET image sequences for exploring early breast cancer diagnosis. Quantitative evaluations, as discussed and validated by clinicians, confirmed the efficiency of the modeling and the usefulness of the new empiric parameter $K_{FPQ}$ to predict tumor glucose metabolism for early uptake. This could be considered as a significant indication for quantification as well as evaluation of early relapse and disease progression during the therapy.

Further research concerning our proposed approach perspectives could be involved regarding extension of the data set images both for micro tumors’ dimension as well as for tumours’ demultiplication.

References


Détermination automatisée des volumes fonctionnels en imagerie d’émission pour les applications en oncologie, 2008 (Thèse de doctorat), INSERM U650 – Laboratoire de Traitement de l’Information Médicale (LaTIM), soutenue le 3 décembre.