POLITECNICO DI TORINO
Repository ISTITUZIONALE

Automatic liver detection and standardised uptake value evaluation in whole-body Positron Emission Tomography/Computed Tomography scans

(Article begins on next page)

19 April 2024

Automated Evaluation of the Liver Standardised Uptake Value in whole-body Positron Emission Tomography/Computed Tomography scans

Stephane Chauvie[a], Elisa Bertone[a], Fabrizio Bergesio[a], Emanuele Roberto[a], Alessandra Terulla[a], Davide Botto[b], Piergiorgio Cerello[c]

[a]*Medical Physics Unit, Santa Croce e Carle Hospital, Cuneo, Italy*
[b]*Dipartimento di Scienza Applicata e Tecnologia, Politecnico di Torino, Torino, Italy*
[c]*INFN, Torino, Italy*

Abstract

Background and Objective: Standardized Uptake Value (SUV), in clinical research and practice, is a marker of tumor avidity in Positron Emission Tomography/Computed Tomography (PET/CT). Since many technical, physical and physiological factors affect the SUV absolute measurement, the liver uptake is often used as reference value both in quantitative and semi-quantitative evaluation. The purpose of this investigation was to automatically detect the liver position in whole-body PET/CT scans and extract its SUV value.
Methods: we developed an algorithm, called LIver DEtection Algorithm (LIDEA), that analyze PET/CT scans, and under the assumption that the liver is a large homogenous area near the centre of mass of the patient, find its position and automatically places a region of interest (ROI) in the liver. This ROI is used to calculates SUV. The algorithm was tested on a population of 630 PET/CT scans coming from more than 60 different scanners. SUV was also manually calculated manually placing a large ROI in the liver, far away form liver's edges.
Results: LIDEA identified the liver with a 97.3 % sensitivity with PET/CT images only and reached a 98.9 % correct detection rate when using the co-registered CT scan to avoid liver misidentification in the right lung.
The average liver SUV obtained with LIDEA was successfully validated against a manual assessment, with no systematic difference (0.11 ± 0.36 SUV units) and a $R^2 = 0.89$ correlation coefficient.
Conclusions: LIDEA was a reliable tool to automatically identify the liver in oncological whole-body PET/CT scans and extract its SUV.

Keywords: Positron Emission Tomography, SUV, Liver.

1. Introduction

Positron Emission Tomography/Computed Tomography (PET/CT) is a fundamental tool in oncology, widely used in the staging, re-staging and follow-up

of several malignant pathologies [1, 2].

Several studies [3] prospect the evaluation of PET/CT scans beside the conventional visual assessment, by using quantitative and semi-quantitative tools. A standard visual assessment is based on the analysis of volumes of tracer hyper-concentration with respect to the surrounding background of healthy tissue. Its principal limitation is the requirement that the background to whom the tracer uptake is compared be constant and independent of the patient physiological state[4]. This often is not the case, both because the patient could lack physiological condition and because several protocol-related factors affect the tracer uptake. ^{18}F-Fluorodeoxyglucose (^{18}F-FDG) measured uptake, for example, depends on a variety of factors related to the scanning procedure and the actual scanner used for image acquisition[5].

The liver uptake in Standardised Uptake Value (SUV) units has been proposed as a reference for PET/CT scan evaluation in different clinical settings [6], both in a qualitative (e.g., uptake of the lesion higher than that of the liver) and semi-quantitative way (e.g. lesion uptake higher than twice the liver uptake)[3]. Even though different independent factors such as Body Mass Index (BMI) and Blood Glucose Level (BGL) [7] influence the liver ^{18}F-FDG uptake, the average liver SUV remains nearly constant (within 5% of the peak value) if the time delay between the tracer injection and the PET/CT acquisition is in the 50-110 min range, with a peak at about 75-80 minutes [8].

In this work we describe and validate a fully automated approach for the liver uptake measurement in whole-body FDG PET/CT scans, which could enable the reproducible calculation of the liver tracer uptake, become a powerful tool for the tumour-to-reference tissue ratio measurement in multi-center clinical trials and be adopted for intra- and inter-patient comparison in clinical applications.

The method would be particularly useful in clinical research applications where the physiological variability associated with direct SUV measurements is unacceptably high and a reference value is necessary.

2. Materials and Methods

2.1. Algorithm Description

The LIver DEtection Algorithm (LIDEA), described in detail in the following, was conceived and structured with the goal of identifying the liver position and evaluating the average liver SUV in a PET/CT scan.

In order to find a large and homogeneous volume inside the 3D scan of the patient, that is a liver candidate, the following steps are taken:

1. masking the voxels outside the patient body;
2. projecting the 3D image on a single 2D coronal image, plane (x, z), of the patient;
3. downscaling the 2D coronal image to obtain average and standard deviation values in volumes larger than the single voxel;

4. finding the x, y, z position of the minimum of the ratio of the standard deviation to average SUV, so as to select a homogeneous volume with relatively high uptake (i.e., a liver candidate).

Initially, the $3D^{original}SUV_{ijk}$ matrix containing the voxel data from the PET/CT slices is extracted from DICOM images. The SUV ina voxel is defined as:

$$SUV = \frac{[A_{tissue}]w_b}{A_{PET}}$$

where $[A_{tissue}]$ is the tracer activity concentration in the voxel, w_b is the patient body weight and A_{PET} is the total activity injected into the patient evaluated at the acquisition time. All the $3D^{original}SUV_{ijk}$ matrix are roto-translated to have the patient in the Head First Supine position. The first step of the algorithm is the removal of voxels outside the patient body from the $3D^{original}SUV_{ijk}$ matrix, that are identifiable as the exterior to the high gradient region between the patient and the surrounding air. The $3D^{original}SUV_{ijk}$ Jacobian matrix (J_{ijk}) is then computed and voxels with values below a threshold (Th) are masked.

The threshold is defined as the average Jacobian plus 7 times its standard deviation ($Th = <J_{ijk}> + 7\sigma_{ijk}$) over two cubic volumes of 5 cm side in two different positions (anterior and posterior to the patient head). Starting from the most superior axial plane and moving in the cranial-caudal direction, all the J_{ijk} voxels that fail to meet the $J_{ijk} < Th$ condition are set to 0 in the $3D^{original}SUV_{ijk}$ matrix.

The second step is the $3D^{original}SUV_{ijk}$ projection on the coronal plane in 2D matrix, calculated as the standard deviation $SUV_{ik}^{\sigma} = \sigma((3D^{original}SUV_{ijk})_j)$ of the voxels along the projection direction (y-axis).

The obtained 2D matrix is then resampled with a pitch (S) so as to obtain the 2D matrices in (x, z) plane of the average (H_{SUV}^{μ}) and standard deviation (H_{SUV}^{σ}). The new matrices are composed of squared pixels of size S. From the $R_{matrix} = H_{SUV}^{\sigma}/H_{SUV}^{\mu}$ ratio, the $i'k'$ minimum position and its $R_{i'k'}$ distance from the Center of Mass (CoM) are determined.

If the $R_{i'k'}$ distance is higher than a reference value or if the $i'k'$ minimum position is located on the right of the CoM (body left), the $i'k'$ point is rejected. The algorithm iterates the search for the minimum until the conditions are satisfied. This step rejects points too far from the CoM (e.g., legs) and in the left part of the body (e.g. heart).

Once the i, k position of the minimum in the coronal projection is obtained, the j depth is recovered by analysing the $3D_{SUV}^{original}$ voxels in the $i'k'$ position and evaluating the minimum of the ratio of the standard deviation and the average value of SUV in the anterior posterior direction, calculated from cubic volumes with side equal to the downscaled pitch S.

The (ijk) position of the minimum is the Reference Liver Position (RLP) as found by the algorithm. The original PET/CT images are tagged with a sphere of 5 cm diameter centered in RLP and the corresponding SUV average and SUV standard deviation are calculated.
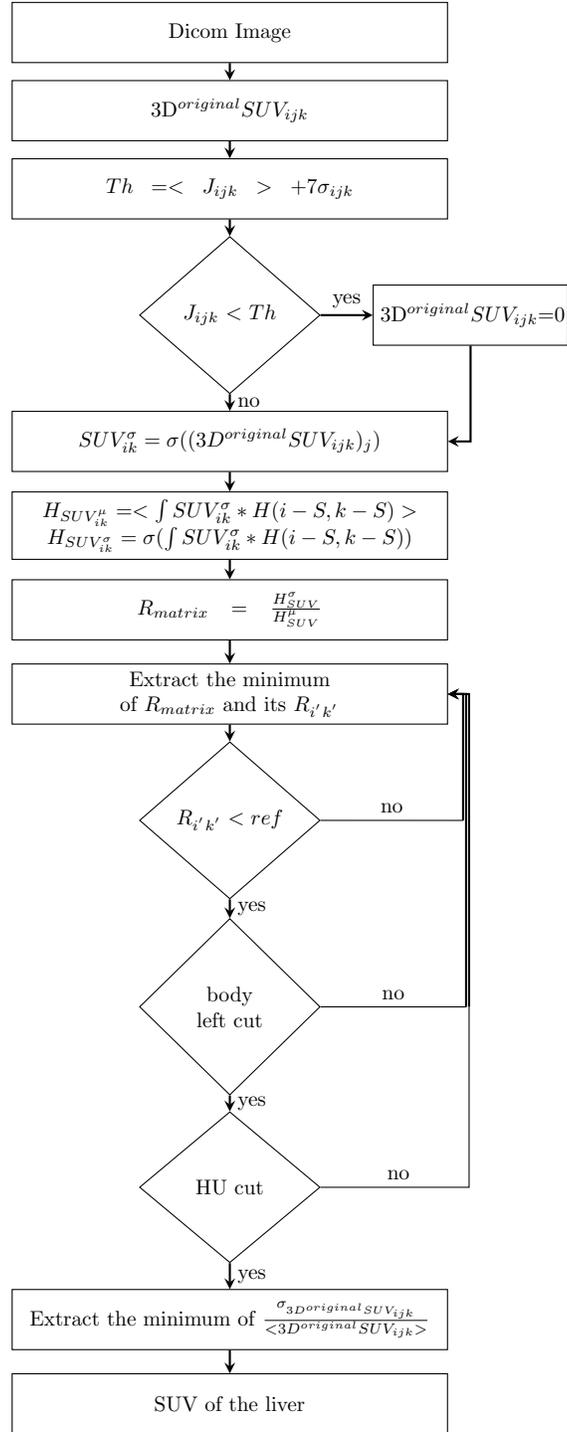
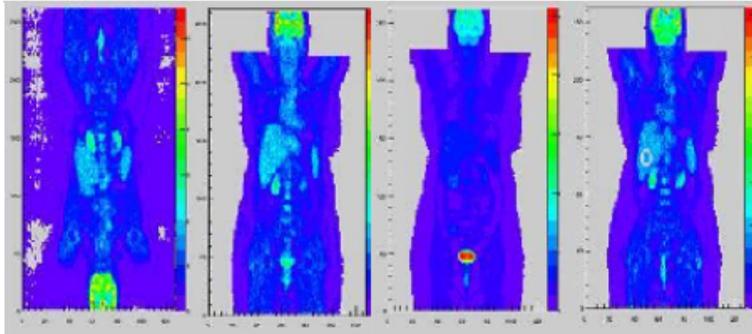Figure 1: Flow diagram of the algorithm

Figure 2: From left to right: coronal slice with the background, coronal slice without the background, projection of $3D^{original}SUV_{ijk}$ matrix standard deviation, coronal slice with the ROI into the liver

2.2. PET/CT scans Dataset

PET/CT scans of patients, enrolled in several multicentric clinical trials, that allowed their use by signing an informed consent were collected in the Cuneo Core Lab [9]. No clinical or personal information were stored in the scan metadata. All the PET/CT scans were acquired according to local site protocols. We considered a cohort with 49.5% of females and 50.5 % of males. The following conditions were required for the inclusion of scans in the Dataset:

- use of a full-ring PET/CT scanner;

- iterative reconstruction applied to PET/CT images;

- availability of CT and attenuation-corrected PET;

- whole body PET/CT scans defined as scans including the region from the base of the skull to the mid-thigh.

A total of 65 different scanners were introduced in this investigation from different vendors and with different characteristic as shown in table 1

Table 1: Characteristics of the different PET/CT scanners used in this investigation

Manufacturers		
	Siemens	10 %
	Philips	23 %
	GE	67 %
Pixel dimensions		
	less than 1 mm	22 %
	range (1 - 1.5) mm	75 %
	more than 1.5 mm	3 %
Slice thickness		
	less than 3 mm	2%
	range (3-4)mm	10%
	range (4-5)mm	54%
	more than 5 mm	34 %

The DICOM headers of PET/CT scans were analysed to extract the following information: body weight (w_b) and height (h_b), injection time and activity ($A_{injected}$), acquisition time. The uptake time (t_{uptake}) was calculated as the difference between the acquisition and tracer administration (injection) time. Wrong parameters in the DICOM headers were not corrected (e.g. wrong injected activity, wrong time, etc...). This was done to test the algorithm in the worst condition, because it should be used in true clinical condition.

All the PET/CT scans were also independently analysed by an expert medical physicist to retrieve the average liver SUV, manually measured in a large Region Of Interest (ROI), with 5 cm diameter, placed in a central position of the liver volume, far away (at least 3 cm) from the liver edge so as to avoid spilling from adjacent structures (such as the gallbladder). The maximum SUV value was the highest voxel value in the ROI, while the average value and the standard deviation were evaluated on all the ROI voxels.

All the images generated by the LIDEA algorithm were checked to verify that the tagged spheres around the RLP were fully positioned inside the liver, following the same rule used to draw manual ROIs. The True Findings Rate (TFR) was defined as the ratio of the spheres completely located in the liver to the total number of analysed PET/CT scans. The False Findings Rate (FFR) was then FFR = 1-TFR.

2.3. Statistical Analysis

In order to test its robustness and verify the possible dependence of its performance on the parameters used within, LIDEA was initially applied to a small cohort of patients (30). The downscaling pitch was set to 1.25 cm, 2.5

6

cm, 5.0 cm and 7.5 cm, a compromise between constraints set by the algorithm speed and the liver size.

Confidence levels of TFR using different sets of parameters were calculated with a p=0.05 agreement level using a binomial distribution.

The best set of parameters was selected with the χ^2 test and used for the analysis of the whole dataset.

Univariate analysis was also carried out to exclude the probability that some variable could influence the application method.

In the full dataset analysis, the correlation between the automated and manual methods was estimated with a linear regression model:

$$SUV_{auto} = p_0 + p_1 SUV_{manual}$$

where p_0 and p_1 are the intercept and slope coefficient, respectively.

Bland-Altman statistics was used to estimate systematic differences between the automated and manual measurements. All the statistical tests were carried out with R [10]. Shapiro and Kolmogorov-Smirnov tests were used to analyse the distribution of SUV values.

3. Results

3.1. Algorithm testing on small cohort

LIDEA was initially applied to a dataset of 30 randomly selected PET/CT scans. T-Students demonstrated no difference in all the parameters between this sample of scans and the whole population of scans. Table 2 shows the liver TFR with different values of the downscaling pitch S. The S=2.5 cm pitch

Table 2: TFR for different values of the downscaling pitch.

S[cm]	TFR	confidence level
1.25	0.966	(0.901,1.031)
2.5	0.966	(0.901,1.031)
5.0	0.897	(0.878,1.096)
7.5	0.400	(0.225,0.575)

provided the best agreement with the manual method and was selected for further analysis. With this value the RLP (table 3) was found inside the liver in all but one cases.

When applying a cut of $32cm$ from the com, that is at 7 standard deviation from average RLP poisition, the leg false findings were easily rejected (table 3) and the second minimum found by the algorithm was correctly positioned in the liver.

Table 3: TFR for liver detection and for different anatomical regions on the small cohort.

	TFR		FFR			
	liver	heart	right lung	left lung	legs	others
no cut	0.966	0.000	0.000	0.000	0.034	0.000
distance cut	1.000	0.000	0.000	0.000	0.000	0.000

Table 4: PET/CT scan parameters of the full dataset as extracted from DICOM header.

	Average	St.Dev.	Median	Min	Max
Body weight[kg]	70	16	68	38	160
Patient height[m]	1.70	0.09	1.70	1.50	1.93
Body mass index[kg/m^2]	23.8	4.0	23.1	15.2	41.5
Injected activity[MBq]	326	94	318	150	756
Uptake time[min]	81.18	39.13	75	30	422
Acquisition Duration (s)	185	55	180	40	300

3.2. Validation on a large cohort

LIDEA was then applied to a set of 630 PET/CT scans, whose summary information is shown in table 4.

Univariate analysis demonstrated that no correlation existed among the characteristics of the patient (weight, height, body mass index and liver SUV), the injected activity and the uptake time, confirming that no population bias was introduced.

The TFR on the large cohort were: 0.93 with no cuts, 0.95 with the distance cut (which eliminates RLPs in the legs) and 0.97 when introducing the boby left cut, as summarised in table 5. A total of 9 right lungs and 8 other structures were still incorrectly detected by the algorithm. In these 9 cases the average and standard deviation SUV values were: (0.63 ± 0.12) and the Hounsfield Unit (HU) of the same cases were: (-692 ± 105). Hence, adding information on the CT, cutting the HU lower than the HU of water, the TFR value raised to 0.989. Seven cases, having an average and standard deviation SUV values of (0.75 ± 0.19), were still misclassified and the RLP was found in the following positio: axillary cavity (2 cases), muscles (4 cases), abdomen (1 case). The corresponding HU were (52 ± 29), not allowing a clean CT-based rejection.

Table 5: TFR for liver detection and for different anatomical regions on the whole population.

	TFR			FFR		
	liver	heart	right lung	left lung	legs	others
original	0.932	0.000	0.000	0.000	0.068	0.000
distance cut	0.954	0.007	0.018	0.009	0.07	0.05
body left cut	0.973	0.000	0.017	0.000	0.000	0.010
HU cut	0.989	0.000	0.000	0.000	0.000	0.011

The liver average SUV for the dataset in the central 5 cm sphere (excluding the 7 false findings) using the manual method was $\mu_{Manual} = (1.90 \pm 0.51)$, while with the automated method it was $\mu_{Auto} = (2.01 \pm 0.55)$. The average value of standard deviation in the sphere was (0.32 ± 0.13) for the automated method.
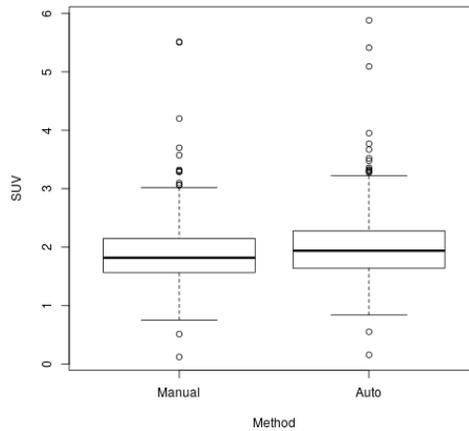

Figure 3: Box plot of SUV extracted automatically with LIDEA (right) and manually (left) from 623 PET/CT scans. The line in the box represents the median, the borders of the box are the first and third quartiles respectively, the horizontal bars are the minimum and maximum values when excluding the outliers.

The distributions of liver average SUV values found automatically by LIDEA and manually by the operator are shown in fig.4.
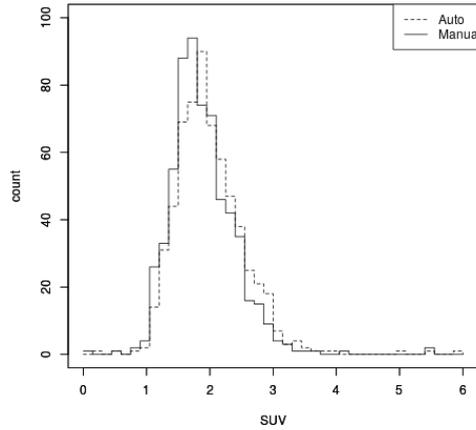

Figure 4: Liver average SUV distribution extracted automatically with LIDEA (dashed line) and manually (full line) from the large cohort of 623 PET/CT scans.

When analysed with the Shapiro-Wilk test, the distributions of the manual and automated liver SUV values are not compatible with a normal distribution. The shapes of the SUV distributions obtained with the two methods were comparable, as shown by the Kolmogorov-Smirnov test (D=0.075, p-value=0.9999). The correlation between the automated and the manual method (fig. 5), when applying a linear regression model, provides the following fit parameter values: $p_0 = 0.08 \pm 0.03$, $p_1 = 1.02 \pm 0.01$, with a correlation coefficient $R^2 = 0.89$.
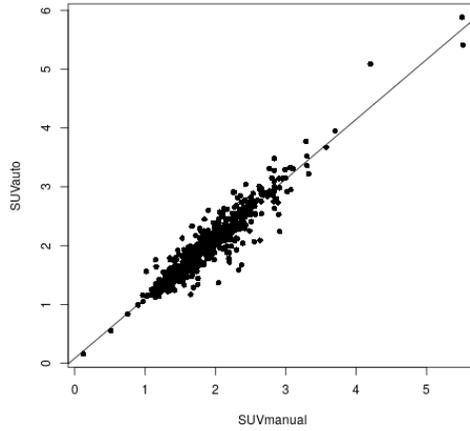
Figure 5: Correlation between the manual and automated average SUV value in the liver.

In spite of a slight bias towards smaller SUV values with the manual method, the correlation between the two methods is satisfactory. In the Bland-Altman plot (fig. 6) the observed systematic difference between the methods $((-0.11 \pm 0.36)$ SUV units) is compatible with zero and much smaller than the liver SUV intrinsic variability, since the average of the standard deviation of the pixel value in a 5 cm diameter ROI inside the liver was (0.32 ± 0.13).
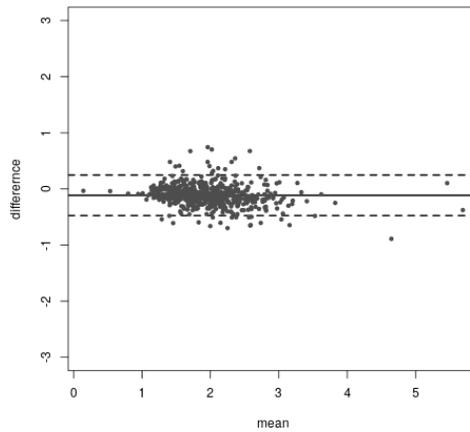

Figure 6: Bland-Altman plot of the difference between automatic and manual average SUV value in the liver, dotted lines are at 95% confidence level.

4. Discussion

The automated segmentation and/or uptake measurement of structures and organs in full-body FDG PET/CT is not extensively addressed in the literature.

Li et al.[11] presented a method to segment the liver in low-contrast CT, based on a segmentation of the liver in PET/CT datasets. The liver segmentation was based on a fixed threshold of SUV > 1.5 to segment liver from other abdomen structures but the algorithm was not applied to whole-body scan but only to 35 normal liver studies coming from a single PET/CT scanner. Moreover no extraction of SUV was done within this investigation.

Bauer et al.[12] performed liver segmentation and an improved initial estimation of the liver SUV on a set of 134 scan from a single PET/CT scanner. In their work they demonstrated a correlation between manual and automatic measurement. The approach used by the author was specifically tailored to the imaging of patients with arms down, and the liver region was automatically identified in subparts of the image relative to the brain, that was always present. In their implementation the authors used a fixed SUV threshold for liver and suggest that additional processing is required to handle major errors in SUV calculation or in quality control applications, where PET image values can be off by a large margin.

Wang et al. [13] developed an automatic anatomy recognition methodology already used in contrast-enhanced diagnostic CT and MR images to automatically detect several organs in body torso (from lung apex to pelvis) PET/CT images. The authors employed 16 whole-body PET/CT images obtained on the same scanner for testing their methods. The average localization error for the liver center was reported as 15 mm on PET/CT scans. No information was given on SUV and intrascanner and interscanner variations was not addressed in this paper.

Hirata et al. [14] proposed a semi-automated algorithm to objectively determine SUV. In this case the nuclear medicine physicians manually place a large spherical volume of interest to roughly enclose the right lobe of the liver. Then the algorithm automatically calculate the 30-mm spherical VOI that has the lowest variability.

Hsu et. al. in their pioneer work proposed a new active contour model, called Poisson Gradient Vector Flow, with genetic algorithm to constructs a scheme to automatically find the contour of liver in the PET images. The contour were found on the 3 image data sets with 16 abdomen PET slice images were used for testing.

At our knowledge this work is the first that prove, that liver could be automatically found and SUV evaluated in a large series of patients acquired in a variety of scanners in many PET/CT site arround the word in normal clinical condition.

In this work, the automated identification of the liver position and the measurement of its average SUV were based on the assumptions that the liver:

1. be large, usually ing 1-1.5 kg;

2. be homogeneous, without anatomical and physiological sub-structures within: indeed, the liver parenchyma is constituted of hepatocytes that are gathered in tubular units with arterial vessels that supply blood and a venous system that eliminates the waste products;

3. be close to the CoM of the patient and mostly in the right part of the body;

4. be more dense than water (i.e., its Hounsfield Units values in the co-registered CT be larger than that of water);

Applying all these conditions, the liver was identified with a 98.9% sensitivity in a large cohort of patients (630 PET/CT scans).
The automated algorithm also shows a remarkable agreement with the independent measurement of average liver SUV obtained from the manual analysis, with no significant systematic difference (0.11 ± 0.36 SUV units).

These results were obtained on a sample of PET/CT scans that did not include strict quality requirements. This was explicitly done to analyze a real-life clinical environment. Hence acquisition parameters, as registered in the DICOM header, could be much smaller or larger than expected, typically because of violations of the required harmonized common protocol [5] or because of wrong insertion of the parameters in the PET/CT scanners done by the technicians. LIDEA indeed proved to reliably find the liver and extract its SUV.

The automatic extraction of SUV from the liver is a useful tool in several applications. It could be used as an image quality control, being SUV values too low or too high representative of some problem in the acquisition. It could be used as a starting point for segmentation algorithm that use liver threshold for identifying lesions. It could be used to automatically calculate indexes such as qPET [15] that is the ratio of lesion $SUV_M AX$ respect to liver SUV. LIDEA permits to perform this action automatically and hence speed-up the referral by imaging expert of oncology PET/CT scans.

5. Conclusion

LIDEA proved to be a reliable algorithm to identify the liver and extract automatically its SUV on a large PET/CT dataset representative of real-life, oncological, whole-body PET/CT scans. These features could be used to automatically assess image quality, as well as a basis for automated PET/CT (semi-)quantitative evaluation.

6. Acknowledgements

Bibliography

[1] A. Jerusalem, R. Hustinx, Y. Beguin, G. Fillet, The value of positron emission tomography (pet) imaging in disease staging and therapy assessment., Ann. Oncol. (13 Suppl 4) (2002) 227–34.

[2] M. Juweid, B. Cheson, Positron-emission tomography and assessment of cancer therapy., N. Engl. J. Med. 354(5), 496-507.

[3] L. Kostakoglu, S. Chauvie, Pet-derived metabolic volume metrics in lymphoma, Clin Transl Imaging (3) (2015) 331–341.

[4] P. E. Kinahan, J. W. Fletcher, Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy, Elsevier 31(6) (2010) 496–505.

[5] R. Boellaard, M. J. ODoherty, W. A. Weber, F. M. Mottaghy, M. N. Lonsdale, S. G. Stroobants, W. J. G. Oyen, J. Kotzerke, O. S. Hoekstra, J. Pruim, P. K. Marsden, K. Tatsch, C. J. Hoekstra, E. P. Visser, B. Arends, F. J. Verzijlbergen, J. M. Zijlstra, E. F. I. Comans, A. A. Lammertsma, A. M. Paans, A. T. Willemsen, T. Beyer, A. Bockisch, C. Schaefer-Prokop, D. Delbeke, R. P. Baum, A. Chiti, B. J. Krause, Fdg pet and pet/ct: Eanm procedure guidelines for tumour pet imaging: version 1.0., Eur J Nucl Med Mol Imaging 37(1) (2010) 181–200.

[6] M. Hutchings, S. Barrington, Pet/ct for therapy response assessment in lymphoma., J. Nucl. Med. (50 Suppl 1) (2009) 21S–30S.

[7] Kamimura, Associations between liver (18)f fluoro-2-deoxy-d-glucose accumulation and various clinical parameters in a japanese population: influence of the metabolic syndrome., ANM 24(3) (2010) 157–61.

[8] E. Laffon, X. Adhoute, H. de Clermont, R. Marthan, Is liver suv stable over time in f-fdg pet imaging?, JNMT 39(4) (2011) 258–63.

[9] S. Chauvie, A. Biggi, A. Stancu, P. Cerello, A. Cavallo, F. Fallanca, F. Ficola, M. Gregianin, U. Guerra, A. Chiaravalloti, O. Schillaci, A. Gallamini, Widen: A tool for medical image management in multicenter clinical trials, Clinical Trials (11(3)) (2014) 355–361.

[10] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (2008).
URL http://www.R-project.org

[11] C. Li, X. Wang, Y. Xia, S. Eberl, Y. Yin, , D. D. Feng, Automated pet-guided liver segmentation from low-contrast ct volumes using probabilistic atlas, Comput. Methods Programs Biomed (2011) 3565–3568.

[12] C. Bauer, S. Sun, W. Sun, J. Otis, A. Wallace, B. J. Smith, J. J. Sunderland, M. M. Graham, M. Sonka, J. M. Buatti, R. R. Beichel, Automated measurement of uptake in cerebellum, liver, and aortic arch in full-body fdg pet/ct scans, Med. Phys. 39(6) (2012) 3112–23.

[13] H. Wang, J. K. Udupa, D. Odhner, Y. Tong, L. Zhao, D. A. Torigian, Automatic anatomy recognition in whole-body pet/ct images, Medical Phy 43 (1) (2016) 613.

[14] K. Hirata, K. Kobayashi, K.-P. Wong, O. Manabe, A. Surmak, N. Tamaki, S.-C. Huang, A semi-automated technique determining the liver standardized uptake value reference for tumor delineation in fdg pet-ct, PLOS ON 9 (8).

[15] S. F. Barrington, R. Kluge, Fdg pet for therapy monitoring in hodgkin and non-hodgkin lymphomas, Eur J Nucl Med Mol Imaging (DOI 10.1007/s00259-017-3690-8) (2017) [Epub ahead of print].

EB, PC and SC are owners of a patent relative to the algorithm.