Master intership report

Developpment of a new diagnosis tool in nuclear medicine

Marie Suau

ABSTRACT

Report of master 2 “ingéniérie de la santé” intership- April to august 2018

Summary

Master internship report 3
I- General Introduction 3
II- Synthesis of the scientific literature 4
a- Neuro-degenerative diseases 4
b- PET scan and neurodegenerative diagnosis 4
c- Disease patterns 5
d- Image registration and segmentation tools 6
e- Iterative reconstruction and statistical variability estimation 7
f- Semi-quantitative methods for automatic classification 7
III- Scientific article 9
IV- Precisions 10
a- Database 10
b- Analysis and complementary results not presented in the article 10
c- Analysis and limitations 12
V- General conclusion 14
VI- Optional appendices 15
Appendix 1: Scenium software 15
Bibliography 17

Master internship report

I- General Introduction

This internship consists in developing and testing a new diagnosis tool for dementia on cerebral PET (positons emission tomography)-scanner with 18-FDG (fluoro-desoxy-glucose).
Neurodegenerative diseases are a recent public health issue, because population aging. Moreover, they cost a lot to society.
Medical treatments are not very efficient, even more when they are taken late in the pathological process. The sooner the disease is diagnosed, the more efficient the treatments are to slow the degenerative process. Treatment procedures are different according to the pathology under consideration. So, early and certain diagnosis is a real medical challenge.
This type of diagnosis is not an easy process. It consists in a clinical approach where complementary exams are not necessary, unless if diagnosis remains uncertain after clinical examination. MRI (magnetic resonance imagery) is the first complementary tool used. Then, if the diagnosis remains not clear, lumbar puncture and nuclear medicine exams are recommended. So, we can see that nowadays TEP scanner does not play the most important role in the diagnostic procedure.
However, it could be useful. Indeed, brain needs glucose. It is its only source of energy. So, the glucose metabolism reflects well the level activity of cortical cells. In cerebral PET scanner, a glucose analogous can be used : it is the 18-FDG. We can consider that it is almost classical glucose, with a radioactive plot: a fluor atom. So, thanks to the 18 FDG PET scan it is possible to reconstruct a 3D map the glucose metabolism in the brain. This map reflects the cerebral activity.
Current images does not perfect reflect the true metabolism. Indeed, positons desintegration is a stochastic process. That is to say that if you want to reconstruct the same patient twice in the same acquisition conditions, you will not obtain the same images. Current reconstruction processes don't take this statistical variability into account. To be closer to reality, one research team designed a new tool to reconstruct PET images, taking into account this variability. This interval-valued reconstruction algorithm NIBEM. For each PET slice, two images are obtained : an inferior and a superior image. Each pixel value is not scalar but interval-valued. It has been shown that NIBEM intervals are >90% statistical confidence intervals.
This kind of tool can be interesting for us to improve the accuracy of neurodegenerative diseases diagnosis. With this additional information, we tried to develop a new semi automatized tool, to help nuclear medicine practitioners diagnose this kind of pathologies earlier.
The project is led by a pluri-disciplinary team : it is a collaboration between a medical team, in Montpellier University hospital Gui de Chauliac, and a computer science team from LIRMM (laboratoire d'informatique, robotique et microélectronique de Montpellier), Montpellier University.

II- Synthesis of the scientific literature

Here we try to resume the principal information needed to understand the paper presented in this report : neuro-degenerative diseases, PET scan used for diagnosis, typical neurodegenerative patterns on PET scan, reconstruction and segmentation tools and semi quantitative methods to make a conclusion.

a- Neuro-degenerative diseases

As the world population ages, neurodegeneration represents a real challenge for both society and medical community. Alzheimer disease, considered as the most common neurodegenerative diseases, affects in terms of prevalence almost 2 - 3% of the population after 65 years old and represents more than 60% of all dementia (Chen, 2018) . It represents a public health problem, in terms of costs, for home care or institutional placements for instance ("World Alzheimer Report 2015, The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends," n.d.). More than financial considerations, it is important to notice that the patient impact is considerable, because of autonomy, functional and cognitive functions decrease. Early and certain diagnosis is a key limitation of current patient management. Indeed, treatments (medicine and most of all patient care) are less efficient with patient state of health worsening. There is currently no way to cure this kind of disease, but its progress can be slow down. Nevertheless, for Alzheimer disease, certain diagnosis is only made post-mortem. A lot of different tools already exist : specific bio-markers, clinical tests, imaging exams (Siderowf et al., 2018) (Bloudek et al., 2011). But the guidelines still only recommend clinical diagnosis. Then, if the conclusion remains uncertain, French scientific societies recommend MRI and lumbar punction. However, some studies have shown that using some semi-quantitative tools can help to obtain better results than clinical diagnosis alone (Minoshima et al., 1995; Perani et al., 2014). So, it shows that development of new diagnosis tools can be useful, in the different kinds of imagery (Habert et al., 2011) (Imabayashi et al., n.d.; the Alzheimer's Disease Neuroimaging Initiative et al., 2018).

b- PET scan and neurodegenerative diagnosis

Cerebral PET (position emission tomography) scintigraphy could be an interesting tool (Tripathi et al., 2014) (Gray et al., 2012)(Rodrigues and Silveira, 2014)(Malpetti et al., 2017), even more than MRI or SPECT (Yuan et al., 2009), because it reflects cerebral glucose metabolism, decreased in some areas in neurodegenerative diseases. But it is not recommended in general practice, because its added value remains contested (Drzezga et al., 2018)(Nestor et al., 2018)(Habert et al., 2014). In theory, impacted territories are different according to the disease under consideration (Fouquet et al., n.d.) (HERHOLZ, 2014). Nevertheless, nuclear images interpretation are still practitioner-dependent because it depends on the experience. To be able to interpret images, we have post treated them.

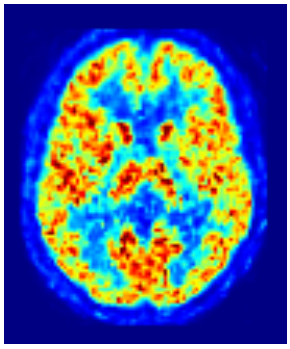


Figure pathological brain, without post treatment

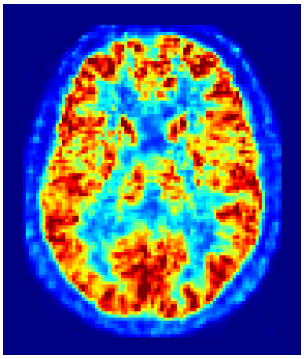
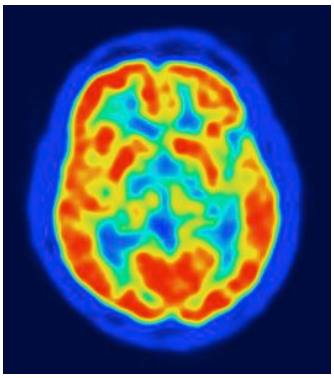


Figure 1 healthy brain, without post treatment

Figure 2 : post treated brain



Even if it improve the visual inspection, this smoothing leads to a loss of information. Thus, visual interpretation alone is not optimal.
In the literature, new softwares have been recently developed to assist doctors in their interpretation (Perani et al., 2014) A new directed interval-based tomographic reconstruction algorithm, called Non-additive Interval Based Expectation Maximization NIBEM could be one of them (Kucharczak et al., 2018). This approach consists in estimating confidence intervals associated to each pixel (ie voxel) the reconstructed image (ie volume) that traduce the uncertainty induced by statistical variations. Here, a possible clinical NIBEM application on cerebral PET images for neurodegenerative diseases diagnosis is investigated.

c- Disease patterns

According to the disease under consideration, the damaged cerebral areas are different (Teune et al., 2010). So, hypometabolism regions are not localized at the same location. On PET scan, some classical patterns are identifiable.

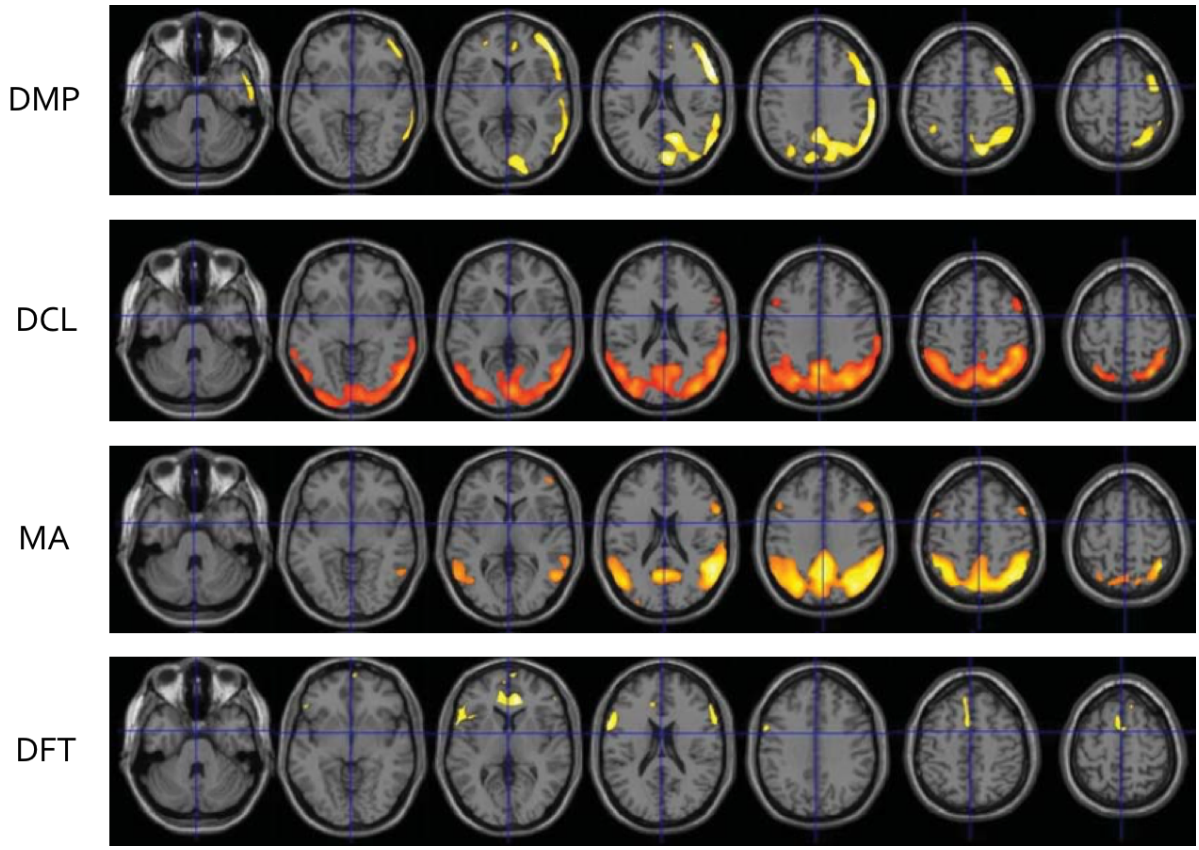


Figure 3 classical patterns for DMP (Parkinson disease) DCL (lewy bodies dementia) MA (Alzheimer disease) and DFT (fronto temporal dementia)

Classical patterns are :

Disease	Alzheimer	Fronto temporal dementia	Lewy bodies dementia	Cortico basal dementia	Primary progressive aphasia	Multi system atrophy	Mild cognitive impairment
ROI	Posterior cyngular gyrus/ temporo parietal cortex/ medial temporal	Frontal lobe/ temporal lobe	Posterior temporal cortex/ occipital lobe/ parieto temporal cortex	Parietal lobe/ fronto parietal cortex	Peri sylvian region	Putamen / cerebellar hemisphere	Posterior cyngular gyrus/ temporo parietal cortex

d- Image registration and segmentation tools

SPM (Statistical Parametric Mapping) is one of the most used tool for statistical analysis in nuclear medicine. It is free, an relatively easy to use. It embedded plenty of image processing tools useful for medical imaging analysis like for instance smoothing, spatial registration on anatomical templates, etc, according to the kind of exam under consideration (MRI, PET scanner ...) . Thanks to image registration, region of interest (ROI) segmentation is made easier. Manual segmentation is fastidious and not

reproducible. We can use an automated tool to segment, using a mask representing the regions of interest.

e- Iterative reconstruction and statistical variability estimation

The reconstruction method used in the following article is the NIBEM algorithm (Kucharczak et al., 2018). It provides intervals instead of single-valued projections. Its specificity relies the fact that, unlike usual reconstruction algorithms that does not have any information about the statistics of the reconstructed values, it directly estimates this information through confidence intervals. The methods currently used for comparing reconstructed activities are mostly empirical, like comparison of SUV described in our experimentation. Associating a confidence interval for each computed value would enable a statistical comparison based on hypothesis testing. Before NIBEM publication, different efforts have been made to try to estimate the unknown statistical variability of the iteratively reconstructed values. Two main approaches have been presented the literature. The first kind of approaches tends to propagate the known variance of measurements to the variance and covariance of reconstructed values. Different methods are derived from this principle: the penalized likelihood maximization(Fessler, n.d.) (Jinyi Qi and Leahy, 2000)(Stayman and Fessler, 2004), the maximum a posteriori algorithm (Li, 2011), the maximum likelihood expectation maximization(Wilson et al., 1994)(Barrett et al., 1994), Bayesian estimates (Higdon et al., 1997)(Sitek, 2012), the origin ensembles (Sitek, 2008) and block iterative methods (Soares et al., 2000)(Soares et al., 2005). They accept the hypothesis that measurements follow a Poisson distribution, that is known to be wrong in practice. The second kind of approach doesn't require any hypothesis on measurements. It uses bootstrap resampling to create replicates (Dahlbom, 2002)(Buvat and Riddell, 2002) (Lartizien et al., 2010). They determine statistical properties of values thanks to this replicates.
Nevertheless, both of this approaches have limitations. They are time consuming, and so calculation is not compatible with clinical use. Moreover, confidence intervals re not trivially with these approaches. Indeed, they give variance or standard deviation. But , we don't know if the distribution is Gaussian on the reconstructed images. So, we can't construct confidence intervals from these information.
The NIBEM approach, described in the paper (Kucharczak et al., 2018), gives a direct estimate of the interval.

f- Semi-quantitative methods for automatic classification

As described in paragraph b, some semi-automatic tools are already existing and have shown their added value. Here we are describing the one already used at the Montpellier University Hospital: the Scenium software (cf appendix 1). It compares the patient exam to be analyzed with a database of healthy patients exams, from patients paired on age with the one investigated. It performs a spatial normalization on an internationally validated template. Then, it segments brain in different areas: the regions of interest ROI (for example right temporal lobe, left temporal lobe, right occipital lobe, left occipital lobe, basal nuclei, right temporo-parietal junction, left temporoparietal junction,…). Then, it performs comparison between the brain of the patient under consideration and the database of registered brains labelled as "normal", for all ROI. Results are presented in the form of standard deviation. An ROI is considered impacted if there is more than 3 standard deviation between the patient and database.

It is not very used on clinical practice.

III- Scientific article

This article is still under redaction. It will be submitted in the European Journal of Nuclear Medicine (EJNM). The research protocol was submitted and accepted by the local ethic committee, the IRB of Montpellier.

RESEARCH

Validation of confidence intervals based reconstruction for differential FDG-PET dementia diagnosis in a clinical context

Florentin Kucharczak[1,2,3*], Marie Suau[1,2,4], Olivier Strauss[1], Denis Mariano-Goulart[2] and Fayçal Ben Bouallegue[2]

Abstract

Background: It has been shown that neuronal activity imaging through ^{18}FDG-PET can help the diagnosis and specification of neurodegenerative dementia etiology. Diagnostic assistance tools has been proposed in the literature making quantitative comparisons to healthy databases. We propose a framework for differential region of interest (ROI) comparison for dementia diagnosis based on confidence interval reconstruction.

Method: Static ^{18}FDG-PET has been performed on 28 patients. Definitive diagnosis was made after enough follow-up for neurologist clinical validation. Using a new framework based on confidence intervals comparison, classification of healthy patients and patients with dementia was performed computing normalized inclusion rate ratio between ROIs.

Results: Results obtained with this tool are better than those obtained with differences indicator. The area under curve is 0.89, against 0.58 for differences indicator.

Conclusion: This preliminary work gave promising results. With a simple test, a patient majority is well classified. Further explorations are needed to validate, with a large scale, multi-centric, prospective study.

Keywords: FDG-TEP; neurodegenerative dementia; confidence interval reconstruction

*Correspondence: florentin.kucharczak@lirmm.fr
[1]LIRMM, Univ. Montpellier, CNRS, Montpellier, France
Full list of author information is available at the end of the article

Background

As the world population ages, neuro-degeneration represents a real challenge for both society and medical community. Alzheimer Disease (AD) is considered as the major neurodegenerative dementia etiology, representing more than 60% of all dementia [1]. In terms of prevalence, its affects 2 - 3% of worldwide population over 65 years old. AD affects patients autonomy, functional abilities and gives rise to cognitive functions decrease. In terms of public health policies, it raises important issues concerning home care and institutional placements of this growing population [2]. The funding of these policies is also becoming a first order problem. Early and certain diagnosis is a key limitation of current patient management. Indeed, treatments (medicine and patient care) are less efficient with worsening patient state of health. Nevertheless, for AD, certain diagnosis is still only available via post-mortem procedures. Important focus have been made to improve specific bio-markers, clinical tests, imaging exams for AD [3] and others dementia etiologies [4]. However, the guidelines still only recommend clinical and Magnetic Resonance Imaging (MRI)-based diagnosis even if some studies have shown that clinical diagnosis could benefit from the development of novel image analysis procedures, for any kind of imagery [5] [6] [7]. Among the different imagery modalities, usual cerebral fluoro-desoxyglucose Positron Emission Tomography (^{18}FDG-PET) has been shown to be an interesting modality for AD diagnosis. Even if its added value against MRI is controversial [8] [9] [10], promising diagnosis results have been obtained with ^{18}FDG-PET [11]. As ^{18}FDG-PET reflects cerebral glucose metabolism and that dementia directly impact metabolism of specific regions of the brain [12] [13], analysis of ^{18}FDG distribution can reach to reliable diagnosis conclusions. However, in addition to being operator dependent, PET interpretation requires post-smoothing of the statistically and spatially noisy reconstructed images. The lost of information induced by this post-procedure step also alters

the hypo-metabolism detection and thus, the ability of the physician to detect early-dementia. These considerations highlight the fact that physicians would benefit from qualitative tools to help them to make their diagnosis [14]. To help the physicians to bring out potential ^{18}FDG hypo-metabolism in specific regions of interest (ROIs) of dementia, some tools were developed. Usually, these methods relies on a first and second order comparison of the ^{18}FDG distribution between the patient considered ROI and an expected distribution computed on a database of healthy patients. The novelty of this work is that the comparison is not performed using any other information than the acquired data itself. Using the recently proposed reconstruction algorithm NIBEM [15] that allows direct estimation of voxel-wise confidence intervals accounting for statistical variability of the acquisitions, this paper presents a new framework for ^{18}FDG-PET differential ROIs comparison. First results of its behavior for early Alzheimer disease diagnosis assistance in a clinical context are presented in this paper. The proposed methodology was also tested on rarer dementia etiologies. First insights are presented in results section.

Methods

Patient characteristics

The main difficulty of the patient recruitment for validation was the necessity to collect raw PET acquisition data. Thus, in this study, 28 patients were retrospectively analyzed. They were classified in two groups: 10 pathological patients and 18 healthy controls. They all underwent a PET exam at the nuclear medicine department in the Gui de Chauliac University Hospital, Montpellier, France, between the November 7th,2016 and the July 31th,2017.The pathological cluster is constituted of 7 women and 3 men. They were between 64 years and 84 years old at the PET exam time. They were labeled pathological, i.e. suffering from Alzheimer disease, according to the most recent medical hypothesis available (from a neurologist, a psychiatrist or geriatrician in most of the case). The clinical diagnosis is considered as gold-standard. They had a medical follow-up for their troubles of more than 3 months. They were indicated to PET exam, among other, because of cognitive, mnesic or dysexecutive disorder. Some physicians already had a clinical hypothesis before scintigraphic conclusion. In the same way, patients with other neurological visible disorder on PET (important stroke consequence for instance), with a different protocol or without clear medical conclusion were also excluded. Patients with a long delay between injection and acquisition were excluded (when superior to 40 minutes). The healthy reference cluster is constituted of 8 women and 10 men. They were between 36 years and 81 years old at the PET exam time. Patients were labeled "normal", that is to say without neurodegenerative process after PET or medical conclusion. Indications for PET exam were varied. We kept healthy controls with a follow up of less than 3 months, when the indication was not mnesic troubles or when there was no doubt on the PET images in the exam conclusion and the probability of neurodegenerative disease was very low before PET exam. All the main characteristics like age, time between injection and data acquisition, glycemia before injection, injected dose and follow up duration are presented in Table 1. Statistical significance of groups characteristics is also computed through Wilcoxon's rank sum test. p-values of the corresponding test are also presented in Table 1.

PET data acquisition

PET exams were performed on a Siemens Biograph mCT 20 Flow PET/CT system (Siemens Medical Solutions Knoxville, USA) with an axial field of view of xxx cm in three-dimensional (3D) mode. The emission data was corrected for attenuation using the corresponding CT scanner. It was also corrected for random, scatter, effects of isotope decay and detector resolution using dedicated manufacturer tools. The static PET scan acquisition was performed 20 to 35 min (healthy controls), 30 to 36 min (AD patients) after injection of 121 to 173 MBq (healthy controls), 121 to 166 MBq (AD patients) of ^{18}FDG.

Data reconstruction

Data acquisition sinograms were reconstructed with an iterative algorithm called NIBEM [15], that allows to reconstruct confidence intervals for each reconstructed voxel. The considered confidence intervals accounts for the uncertainty associated with the statistical variations of positron emission. For each patient, 3D corrected sinograms were reconstructed with NIBEM into 109 image planes of 200x200 voxels; each voxel was 4.1x4.1x2 mm in size. 3D acquisition sinograms were previously resorted into a stack of 2-dimensional (2D) sinograms using the Fourier rebinning algorithm (FORE). Concerning NIBEM setup, 70 iterations were used, in order to match the corresponding iteration level of OSEM algorithm used routinely for this type of clinical applications at our institution. As the algorithm used is interval-valued, the value associated to each voxel i is an interval $\overline{f_i}$, $\overline{f_i}$ and $\underline{f_i}$ respectively accounting for the estimated superior bound and the estimated inferior bound for voxel i. Illustration of the reconstruction procedure is graphically presented in Fig 1 .

	Healthy controls (N = 18)	Alzheimer's disease (N = 10)	p-value
Male	10 (56%)	3 (30%)	0.25
Age (yrs)	63 ± 13 [36 81]	76 ± 6 [64 84]	0.003
Follow-up (days)	188 ± 181[0 590]	320 ± 204 [101 606]	0.14
Injected activity (MBq)	147 ± 16 [121 173]	141 ± 18 [121 166]	0.31
Glycemia (g/L)	1.07 ± 0.24 [0.71 1.62]	0.96 ± 0.07 [0.85 1.10]	0.30
Injection-acquisition delay (min)	30 ± 3 [20 35]	32 ± 3 [30 36]	0.08

Table 1 Characteristics of the study groups. Categorical variables are given as number (percentage), p-value is obtained using Fisher's exact test. Continuous variables are given as mean ± standard deviation [range], p-value is obtained using Wilcoxon's rank sum test.

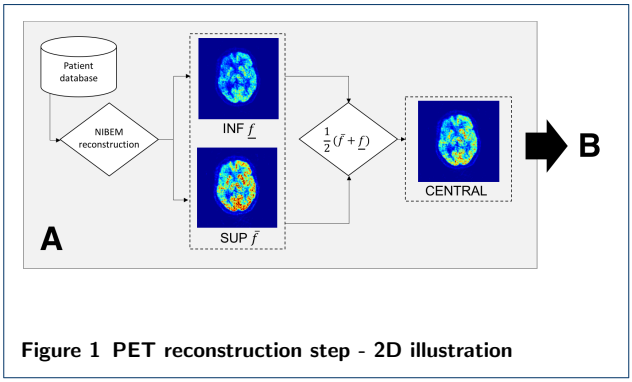


Figure 1 PET reconstruction step - 2D illustration

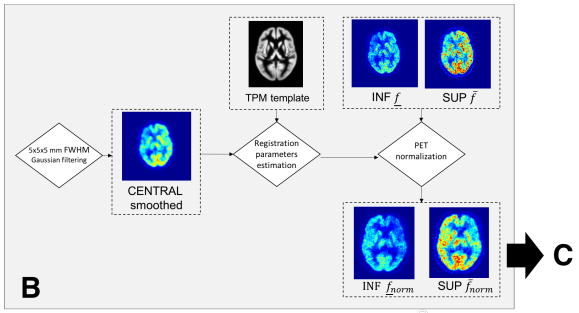


Figure 2 PET normalization step

Interval-valued PET spatial normalization

As the proposed methodology for dementia diagnosis is based on differential comparison of different ROIs FDG concentration, it is mandatory to use a reproducible procedure to segment these ROIs on each patients. To do so, PET normalization was performed using the TPM template proposed on Statistical Parametric Software (SPM)[1]. The spatial normalization parameters where estimated over the central image reconstructed with NIBEM $(0.5 \times (\overline{f_i} + \underline{f_i}))$ smoothed by a 5 mm FWHM Gaussian filter. NIBEM lower and upper bounds images were then spatially normalized with trilinear interpolation using the estimated transformation parameters without other processing. Graphical illustration of the spatial normalization procedure is proposed in Fig 2. .

ROI segmentation

Based on the recommendations of the physicians of our institution and numerous articles of the literature, the focus was made on 4 ROIs of each hemisphere for AD differential diagnosis. The ROIs were the following: right (R) and left (L) frontal lobes, R and L medial temporal regions, R and L temporal lobes, R and L parietal lobes. The atlas used to build our is the neuromorphometrics atlas[2], already spatially normalized to the template we used for PET spatial normalization.

[1]https://www.fil.ion.ucl.ac.uk/spm/
[2]http://www.neuromorphometrics.com/

ROI comparison

In order to distinguish if an interval distribution is different from another, we need to compare two interval-valued distributions. Do do so, the literature doesn't offer a lot of statistical tests generalized to intervals. In this work, we propose to compare the interval-valued distribution through the interval intersection rate criterion. For each interval associated to each of the N voxels of ROI_1, we tested whether it intersects or not with each interval associated to each of the M voxels of ROI_2. The inclusion rate between ROI_1 and ROI_2 that we denote $\mathrm{IR}_{1,2}$ is normalized by the total number of inclusion tests performed.

The intersection function $\mathbf{g}$ between two intervals $[f_1]$ and $[f_2]$ is defined by :

$$\mathbf{g}([f_1],[f_2]) = \begin{cases} 0 & \text{if } \underline{f_1} > \overline{f_2} \text{ or } \overline{f_1} < \underline{f_2}, \\ 1 & \text{else.} \end{cases} \quad (1)$$

$\mathrm{IR}_{1,2}$ is thus defined as:

$$\mathrm{IR}_{1,2} = \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} \mathbf{g}([f_n],[f_m]). \quad (2)$$

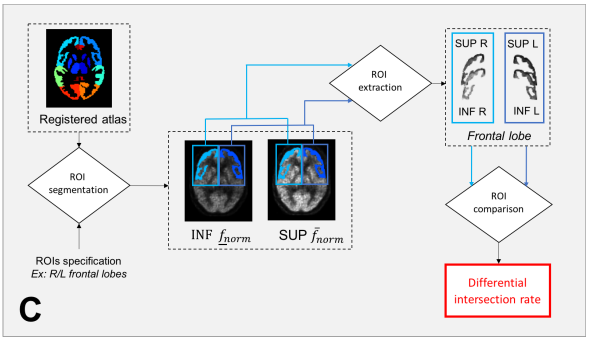


Figure 3 PET segmentation and ROI comparison step

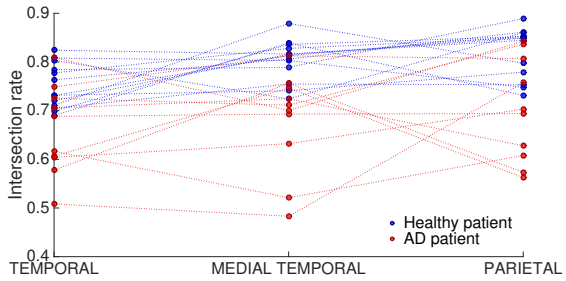


Figure 4 Intersection rate for AD and healthy patients

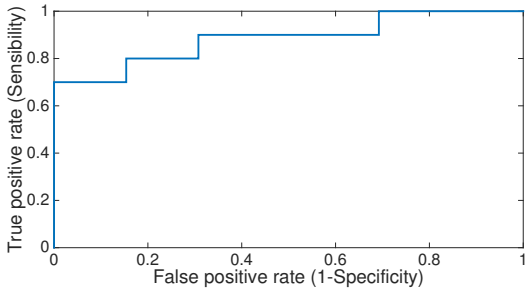


Figure 5 ROC curve of the proposed differential diagnosis test (AUC = 0.89)

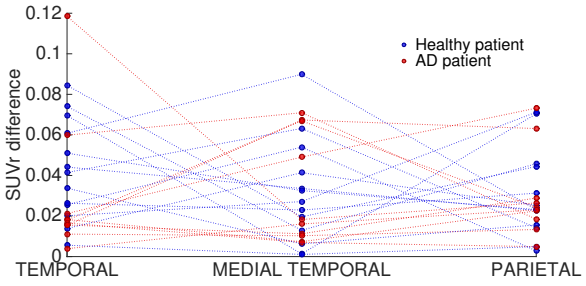


Figure 6 Absolute SUVr difference for AD and healthy patients

$IR_{1,2} = 1$ means that all intervals of both ROIs intersect. At the opposite, $IR_{1,2} = 0$ means that no intervals intersects between two ROIs.

Spatial normalization and ROI segmentation are graphically illustrated in Fig 3 .

Firstly, we computed SUVr method: mean value of each ROI is calculated, using reconstructed values from classical method. Means are normalized with cerebellum mean activity. Then, difference between left and right means are computed. We used same ROI than for intersection rate, that is to say frontal lobe, temporal lobe and temporal medial area. As for intersection rate, ROC curve is drawn.

Results

Our preliminary results are presented on Fig4. Pathologic subjects are represented with red lines and normal subjects with blue lines. Patients are classified with the following criterions: - if one rate (or more) is lower than threshold,the subject is pathological. - if all the intersection rates are superior to the threshold, then subject is classified normal. The corresponding ROC curve is representing on Fig5.The AUC is 0.89. We do the same with difference calculation with SUVr method. Pathologic subjects are represented with red lines and normal subjects with blue lines. Patients are classified with the following criterions: - if one rate (or more) is lower than threshold,the subject is pathological. - if all the intersection rates are superior to the threshold, then subject is classified normal. Results are presented on Fig6. The corresponding ROC curve is representing on Fig7.The AUC is 0.58. So, our new approach gives better results, with an higher AUC.

Conclusion and perspectives

As seen in previous section, this new diagnosis method gives promising results. The AUC is risen from 0.58 to 0.89 between a classical semi-quantitative approach and our innovative approach. With a very basic procedure, we can diagnose a large majority of patients. So, we can imagine that if we add other parameters in our decision algorithm, like Scheltens score (MRI hippocamp atrophy score), MMSE score, ..., these good results can be even greater. Obviously, this work gives preliminary results for further investigations. This conclusion has to be nuanced and some limitations have to pointed out. Our basis is too small. The design is not

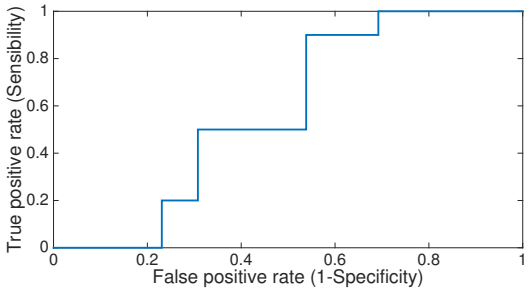


Figure 7 ROC curve corresponding to differential SUVr diagnosis test based on absolute mean comparison (AUC = 0.58)

optimal: it is a retrospective, case control study. The follow up period is not standardized. Some patients had already their diagnosis conclusion before PET examination. A prospective, multi centric and larger clinical trial as to be performed to definitively validate our approach, results presented here are promising. A standardized procedure would help us going farther. Moreover, we could assess our method in comparison with other indicators: MRI conclusion, MMSE score,...

Compliance with ethical standards

Conflict of interest

All authors declare that they have no conflict of interest.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

For this type of (retrospective) study, formal consent is not required.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]LIRMM, Univ. Montpellier, CNRS, Montpellier, France. [2]Department of Nuclear Medicine, Montpellier University Hospital, Univ. Montpellier, , Montpellier, France. [3]Siemens Healthineers, , Saint Denis, France. [4]Mines Saint-Etienne, , Saint-Etienne, France.

References

1. Chen, J.J.: Functional MRI of brain physiology in aging and neurodegenerative diseases. NeuroImage (2018). doi:10.1016/j.neuroimage.2018.05.050. Accessed 2018-06-01
2. Prince, M., Wimo, A., Guerchet, M., Ali, G.-C., Wu, Y.-T., Prina, M., International, A.D.: Worls Alzheimer Report 2015 The global impact of dementia. An analysis of prevalence, incidence, cost and trends. PhD Thesis (2016)
3. Bloudek, L.M., Spackman, D.E., Blankenburg, M., Sullivan, S.D.: Review and Meta-Analysis of Biomarkers and Diagnostic Imaging in Alzheimer's Disease. Journal of Alzheimer's Disease **26**(4), 627–645 (2011). doi:10.3233/JAD-2011-110458. Accessed 2018-06-01
4. Siderowf, A., Aarsland, D., Mollenhauer, B., Goldman, J.G., Ravina, B.: Biomarkers for cognitive impairment in Lewy body disorders: Status and relevance for clinical trials: Biomarkers of Cognitive Impairment. Movement Disorders **33**(4), 528–536 (2018). doi:10.1002/mds.27355. Accessed 2018-06-01
5. Habert, M.-O., Horn, J.-F., Sarazin, M., Lotterie, J.-A., Puel, M., Onen, F., Zanca, M., Portet, F., Touchon, J., Verny, M., Mahieux, F., Giron, A., Fertil, B., Dubois, B.: Brain perfusion SPECT with an automated quantitative tool can identify prodromal Alzheimer's disease among patients with mild cognitive impairment. Neurobiology of Aging **32**(1), 15–23 (2011). doi:10.1016/j.neurobiolaging.2009.01.013. Accessed 2018-06-01
6. Liu, M., Cheng, D., Wang, K., Wang, Y.: Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis. Neuroinformatics (2018). doi:10.1007/s12021-018-9370-4. Accessed 2018-06-01
7. Imabayashi, E., Matsuda, H., Asada, T., Ohnishi, T.: Superiority of 3-dimensional stereotactic surface projection analysis over visual inspection in discrimination of patients with very early Alzheimer's disease from controls using brain perfusion SPECT. The Journal of Nuclear Medicine **45**(9), 1450 (2004)
8. Drzezga, A., Altomare, D., Festari, C., Arbizu, J., Orini, S., Herholz, K., Nestor, P., Agosta, F., Bouwman, F., Nobili, F., Walker, Z., Frisoni, G.B., Boccardi, M.: Diagnostic utility of 18f-Fluorodeoxyglucose positron emission tomography (FDG-PET) in asymptomatic subjects at increased risk for Alzheimer's disease. European Journal of Nuclear Medicine and Molecular Imaging (2018). doi:10.1007/s00259-018-4032-1. Accessed 2018-06-01
9. Nestor, P.J., Altomare, D., Festari, C., Drzezga, A., Rivolta, J., Walker, Z., Bouwman, F., Orini, S., Law, I., Agosta, F., Arbizu, J., Boccardi, M., Nobili, F., Frisoni, G.B.: Clinical utility of FDG-PET for the differential diagnosis among the main forms of dementia. European Journal of Nuclear Medicine and Molecular Imaging (2018). doi:10.1007/s00259-018-4035-y. Accessed 2018-06-01
10. Habert, M.-O., Bera, G., Yeni, N., Kas, A.: Intérêt et place de la scintigraphie cérébrale dans le diagnostic de démence en 2014. Gériatrie et Psychologie Neuropsychiatrie du Vieillissement **12**(4), 413–423 (2014)
11. Yuan, Y., Gu, Z.-X., Wei, W.-S.: Fluorodeoxyglucose–Positron-Emission Tomography, Single-Photon Emission Tomography, and Structural MR Imaging for Prediction of Rapid Conversion to Alzheimer Disease in Patients with Mild Cognitive Impairment: A Meta-Analysis. American Journal of Neuroradiology **30**(2), 404–410 (2009). doi:10.3174/ajnr.A1357. Accessed 2018-06-01
12. Fouquet, M., Villain, N., Chételat, G., Eustache, F., Desgranges, B.: Cerebral imaging and physiopathology of Alzheimer's disease. Psychologie & neuropsychiatrie du vieillissement **5**(4), 269–79 (2007)
13. Herholz, K.: Guidance for reading FDG PET scans in dementia patients. THE QUARTERLY JOURNAL OF NUCLEAR MEDICINE AND MOLECULAR IMAGING **58**(4), 332–343 (2014)
14. Perani, D., Della Rosa, P.A., Cerami, C., Gallivanone, F., Fallanca, F., Vanoli, E.G., Panzacchi, A., Nobili, F., Pappatà, S., Marcone, A., Garibotto, V., Castiglioni, I., Magnani, G., Cappa, S.F., Gianolli, L.: Validation of an optimized SPM procedure for FDG-PET in dementia diagnosis in a clinical setting. NeuroImage: Clinical **6**, 445–454 (2014). doi:10.1016/j.nicl.2014.10.009. Accessed 2018-06-01
15. Kucharczak, F., Loquin, K., Buvat, I., Strauss, O., Mariano-Goulart, D.: Interval-based reconstruction for uncertainty quantification in PET. Physics in Medicine & Biology **63**(3), 035014 (2018). doi:10.1088/1361-6560/aa9ea6. Accessed 2018-06-01

IV- Precisions

a- Database.

To include patient in the database, he has to satisfy some criteria. For the purpose of his study, acquisitions of more than 300 patients (all the patients who underwent a PET scanner at the Montpellier University Hospital, between November 2016 and July 2017, with the "dementia" protocol) were collected. However, after analysis, some of these patients were not addressed for cognitive troubles, but for other indications like Kleine Levin syndrome, dystonia, tremor, …
Then, some patients addressed for the right indication: "cognitive troubles", didn't have a known clinical conclusion until now. That is to say that the neurologist didn't conclude on the underlying pathology. We can't keep this patients, because we want to confront our results with the clinical conclusion.
The third limitation was the follow-up length. For this work, we only want to keep patients with a minimal follow-up of 2 months.
So, considering all this limitations, we kept 10 patients suffering from Alzheimer disease and 18 healthy control patients.

b- Analysis and complementary results not presented in the article

We compared our results with difference indicator. To do so, we computed the mean value of ROIs with the classic reconstructions. Then we normalized this value with mean global cerebellum activity value. We computed right and left differences of this indicator on ROIs (frontal lobe, temporal lobe, medial temporal cortex).

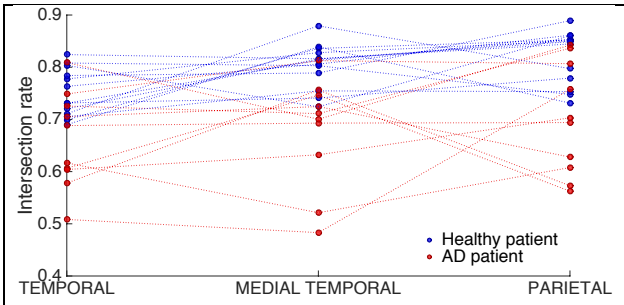


Figure 4 intersection rate on ROIs. Red line: Alzheimer subjects/ Blue line: healthy subjects

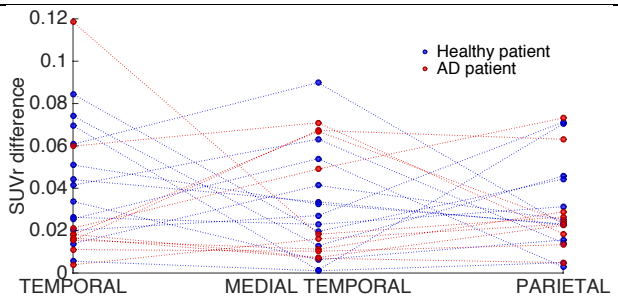


Figure 5 difference rate with SUVr method. Red line: Alzheimer subjects/ Blue line: healthy subjects

Then we classified patients according to the results. A subject is classified as pathological if one indicator is inferior to a pre-defined threshold on one or different ROI. The same approach was used with SUVr criteria and differences computation. The corresponding ROC curves are :

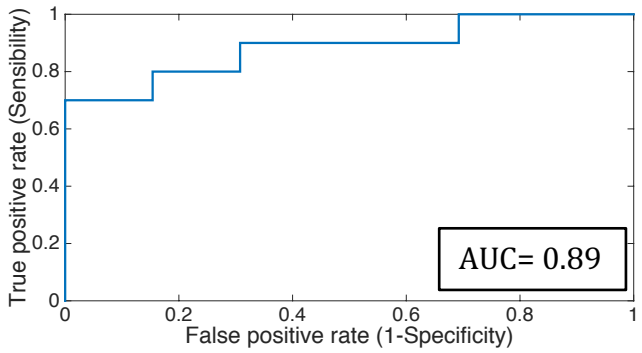


Figure 6 ROC curve with interaction rate

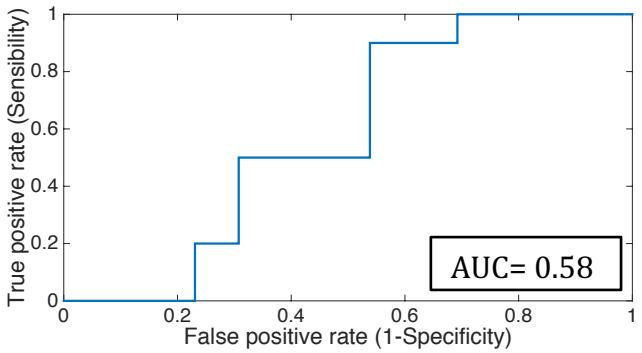


Figure 7 ROC curve with SUVr method and difference computation

We see that the AUC (Area Under the Curve) is lower with SUVr method.
We computed also intersection rates for patients with another disease than Alzheimer disease. We didn't do ROC curve because we have only one subject per disease. But it seems that intersection rate is lower in the ROIs that are characteristics of the considered disease.
We computed also intersection rates different from bilateral rates. We studied rates between all the structures. We see that rates depend on metabolism level. Indeed, rates are higher on healthy subjects. But they also depend on ROI under consideration. We observe same tendencies for healthy and pathological subjects.

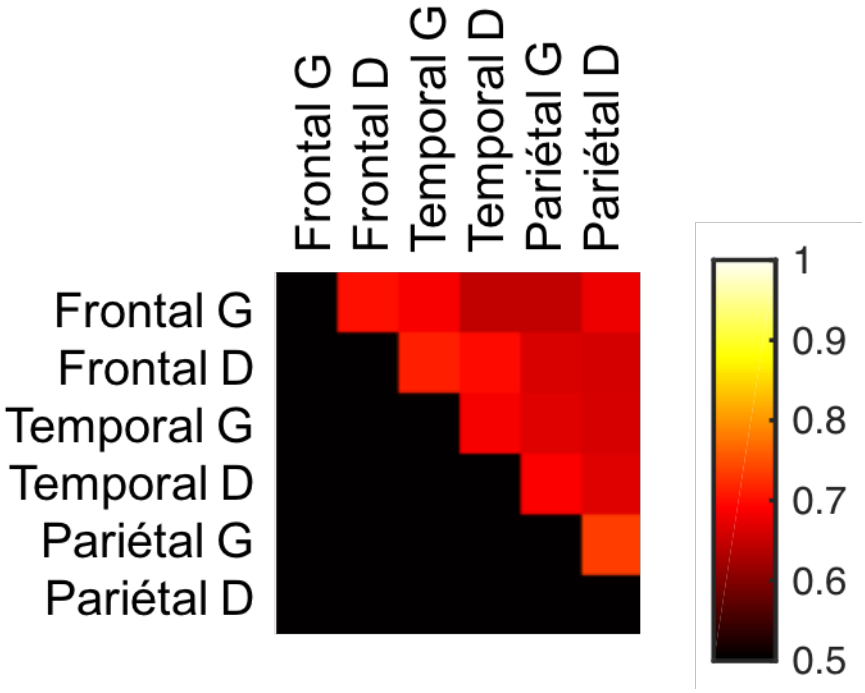


Figure 8 intersection rates on alzheimer subjects.

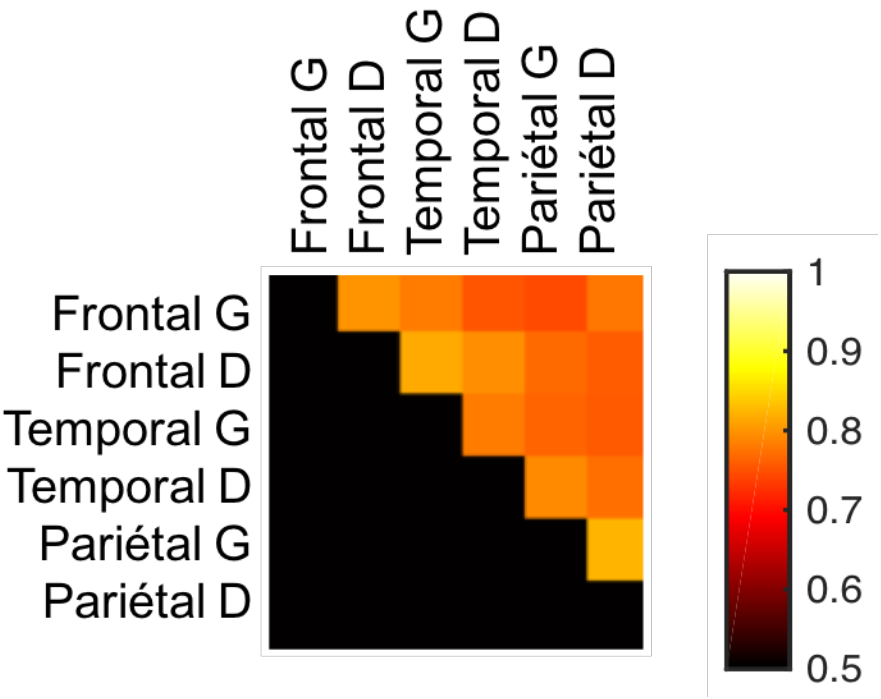


Figure 9 Intersection rates on healthy patients

Another investigation was led. We provided all the exams, including Scenium results, to ax experienced nuclear medicine physician. He was encouraged to state a binary result: the patient is healthy or not, thanks to images and Scenium score. His evaluation gave a sensitivity of 0.9 and a specificity of 0.89.

	Pathologic	Healthy	total
Positive test : classified abnormal	9	2	11
Negative test: classified normal	1	16	17
Total	10	18	

c- Analysis and limitations

Ideally, intersection rates should be lower only in pathological ROI. So, area under curve (AUC from ROC curve) might be closer to 1.
As we can see in the article, results are promising. AUC is clearly higher than 0.5. Moreover, results with the other types of dementias are also interesting. It could be interesting to collect more data of patients with these diseases to statistically validate these insights.
Segmentation plays a central role and a bad segmentation can biased results. Thanks to the matrices presented in the previous section, we can see that our segmentation presents limitations. We have also seen that in ROI with a low signal (superior parietal) the interval are larger. So, because of both, we can hypothesize that intersection rate depends on metabolism, but also on the ROI under consideration.

We can see, thanks to our last test involving nuclear physician for statement, that experience still seems to give better results, and automatic tools are far from replacing doctors.

However, we have to nuance results. There are some limitations. The first weak point is the size of the database. Secondly, sensitivity and specificity are artificially increased by the design, retrospective study, witness/case. To definitively validate this work, we should run a new study: a multicentric prospective clinical trial. We should include subjects following a pre-defined standardized protocol and follow them during a determined duration. And, maybe we could go further and compare our tool with other information than clinical diagnosis. We could for example compare to Scenium results, to MRI conclusion, to MMS score (it's a score used to assess cognitive decline).

V- General conclusion

The presented work is interesting under different viewpoints. First of all, it is one of the first proposition of automatized tool to diagnose patients suspected from Alzheimer disease. Moreover, it illustrates a concrete clinical application for NIBEM reconstruction algorithm.

This work is only preliminary and need to be pursued. Some underlying problems have to be overpassed. Same study should be performed on a prospective way, with an important number of included subjects. We can also imagine that the patients undergo other tests, to compare our method, like MRI and lumbar punction. We could assess the added value of our method. An increased subject number and a prospective design would give a better evidence level.

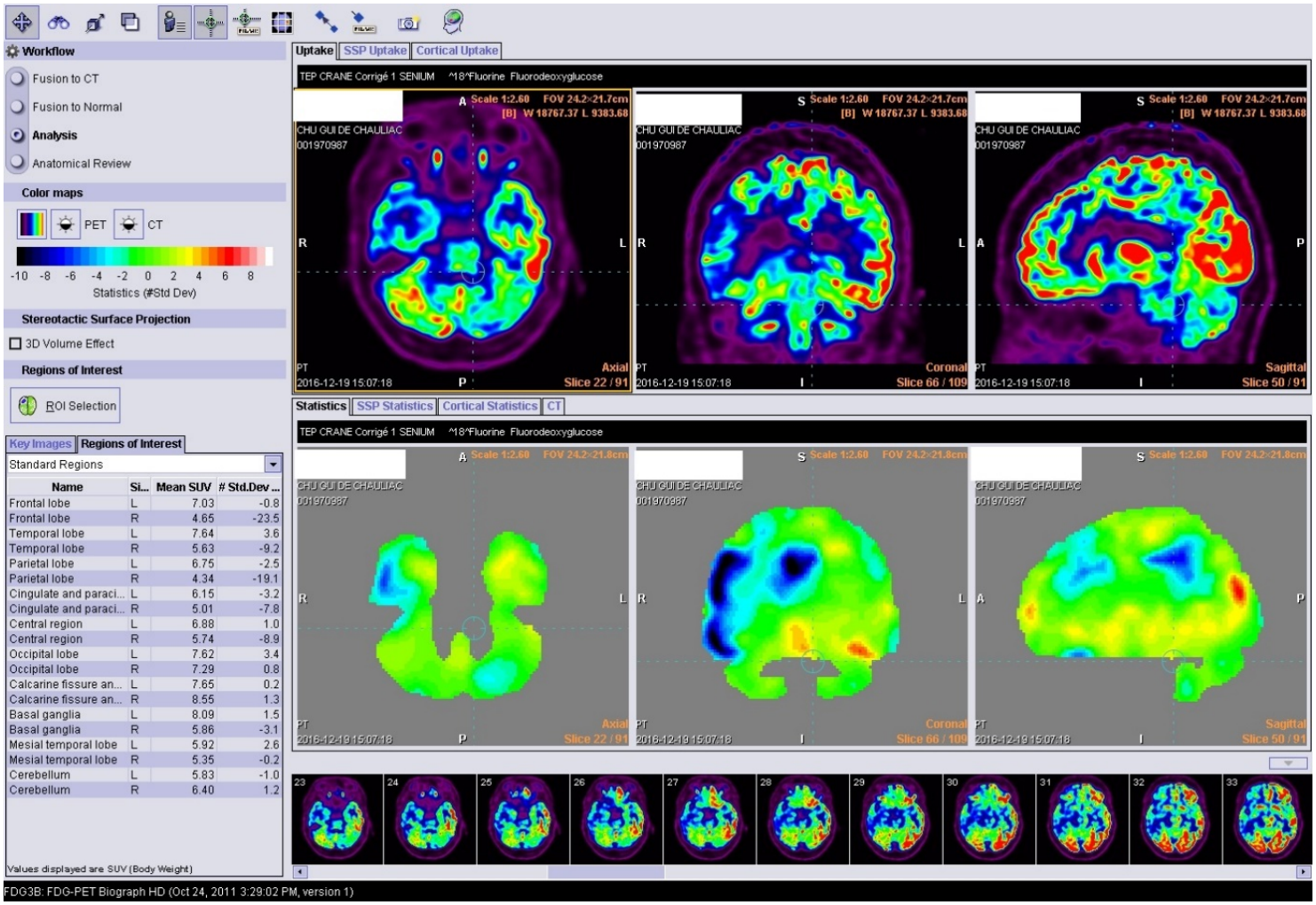
We have shown that, just on images, without other parameters taken into account, we can obtain a valid tool to semi automatically diagnose patients with neurodegenerative patients. We can imagine improving our tool by adding some other parameters like age, comorbidities, MRI Scheltens score,.. It also could be compared with a deep learning approach. But, to perform this kind of comparison, database size is the key limitation point.

This project was very interesting for me. I worked on both engineering and healthcare parts. This topic was on the crossroad of my two domains of competencies. More than interesting, I think that it will be useful for my future career. I perfectionated myself on technical competencies, and I met a lot of important people for continuation.

VI- Optional appendices

Appendix 1: Scenium software

Figure 10 example of Scenium results


Workflow
Fusion to CT
Fusion to Normal
Analysis
Anatomical Review
Color maps
PET
CT
-10 -8 -6 -4 -2 0 2 4 6 8
Statistics (#Std Dev)
Stereotactic Surface Projection
3D Volume Effect
Regions of Interest
ROI Selection
Key Images
Regions of Interest
Standard Regions
Name Si... Mean SUV # Std.Dev ...
Frontal lobe L 7.03 -0.8
Frontal lobe R 4.65 -23.5
Temporal lobe L 7.64 3.6
Temporal lobe R 5.63 -9.2
Parietal lobe L 6.75 -2.5
Parietal lobe R 4.34 -19.1
Cingulate and paraci... L 6.15 -3.2
Cingulate and paraci... R 5.01 -7.8
Central region L 6.88 1.0
Central region R 5.74 -8.9
Occipital lobe L 7.62 3.4
Occipital lobe R 7.29 0.8
Calcarine fissure an... L 7.65 0.2
Calcarine fissure an... R 8.55 1.3
Basal ganglia L 8.09 1.5
Basal ganglia R 5.86 -3.1
Mesial temporal lobe L 5.92 2.6
Mesial temporal lobe R 5.35 -0.2
Cerebellum L 5.83 -1.0
Cerebellum R 6.40 1.2
Values displayed are SUV (Body Weight)
Uptake
SSP Uptake
Cortical Uptake
TEP CRANE Corrigé 1 SENIUM ^18^Fluorine Fluorodeoxyglucose
CHU GUI DE CHAULIAC
001970987
Axial
Coronal
Sagittal
Slice 22 / 91
Slice 66 / 109
Slice 50 / 91
2016-12-19 15:07:18
Statistics
SSP Statistics
Cortical Statistics
CT
23 24 25 26 27 28 29 30 31 32 33
FDG3B: FDG-PET Biograph HD (Oct 24, 2011 3:29:02 PM, version 1)

Bibliography

Barrett, H.H., Wilson, D.W., Tsui, B.M.W., 1994. Phys. Med. Biol. 39, 833–846.

Bloudek, L.M., Spackman, D.E., Blankenburg, M., Sullivan, S.D., 2011. J. Alzheimers Dis. 26, 627–645.

Buvat, I., Riddell, C., 2002. A bootstrap approach for analyzing the statistical properties of SPECT and PET images, in: 2001 IEEE Nuclear Science Symposium Conference Record (Cat. No.01CH37310). Presented at the 2001 IEEE Nuclear Science Symposium Conference Record, IEEE, San Diego, CA, USA, pp. 1419–1423.

Chen, J.J., 2018. NeuroImage.

Dahlbom, M., 2002. IEEE Trans. Nucl. Sci. 49, 2062–2066.

Drzezga, A., Altomare, D., Festari, C., Arbizu, J., Orini, S., Herholz, K., Nestor, P., Agosta, F., Bouwman, F., Nobili, F., Walker, Z., Frisoni, G.B., Boccardi, M., 2018. Eur. J. Nucl. Med. Mol. Imaging 45, 1487–1496.

Fessler, J.A., n.d. 16.

Fouquet, M., Villain, N., Chételat, G., Eustache, F., Desgranges, B., n.d. 31.

Gray, K.R., Wolz, R., Heckemann, R.A., Aljabar, P., Hammers, A., Rueckert, D., 2012. NeuroImage 60, 221–229.

Habert, M.-O., Bera, Gé., Yeni, N., Kas, A., 2014. 12, 11.

Habert, M.-O., Horn, J.-F., Sarazin, M., Lotterie, J.-A., Puel, M., Onen, F., Zanca, M., Portet, F., Touchon, J., Verny, M., Mahieux, F., Giron, A., Fertil, B., Dubois, B., 2011. Neurobiol. Aging 32, 15–23.

HERHOLZ, K., 2014. 58, 12.

Higdon, D.M., Bowsher, J.E., Johnson, V.E., Turkington, T.G., Gilland, D.R., Jaszczak, R.J., 1997. IEEE Trans. Med. Imaging 16, 516–526.

Imabayashi, E., Matsuda, H., Asada, T., Ohnishi, T., Sakamoto, S., Nakano, S., Inoue, T., n.d. 9.

Jinyi Qi, Leahy, R.M., 2000. IEEE Trans. Med. Imaging 19, 493–506.

Kucharczak, F., Loquin, K., Buvat, I., Strauss, O., Mariano-Goulart, D., 2018. Phys. Med. Biol. 63, 035014.

Lartizien, C., Aubin, J.-B., Buvat, I., 2010. IEEE Trans. Med. Imaging 29, 1442–1454.

Li, Y., 2011. Phys. Med. Biol. 56, 1083–1103.

Malpetti, M., Ballarini, T., Presotto, L., Garibotto, V., Tettamanti, M., Perani, D., Alzheimer's Disease Neuroimaging Initiative (ADNI) database;, Network for Efficiency and Standardization of Dementia Diagnosis (NEST-DD) database, 2017. Hum. Brain Mapp. 38, 4212–4227.

Minoshima, S., Frey, K.A., Koeppe, R.A., Foster, N.L., Kuhl, D.E., 1995. 11.

Nestor, P.J., Altomare, D., Festari, C., Drzezga, A., Rivolta, J., Walker, Z., Bouwman, F., Orini, S., Law, I., Agosta, F., Arbizu, J., Boccardi, M., Nobili, F., Frisoni, G.B., 2018. Eur. J. Nucl. Med. Mol. Imaging 45, 1509–1525.

Perani, D., Della Rosa, P.A., Cerami, C., Gallivanone, F., Fallanca, F., Vanoli, E.G., Panzacchi, A., Nobili, F., Pappatà, S., Marcone, A., Garibotto, V., Castiglioni, I., Magnani, G., Cappa, S.F., Gianolli, L., 2014. NeuroImage Clin. 6, 445–454.

Rodrigues, F., Silveira, M., 2014. Longitudinal FDG-PET features for the classification of Alzheimer's disease, in: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Presented at the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Chicago, IL, pp. 1941–1944.

Siderowf, A., Aarsland, D., Mollenhauer, B., Goldman, J.G., Ravina, B., 2018. Mov. Disord. 33, 528–536.
Sitek, A., 2008. Phys. Med. Biol. 53, 3201–3216.
Sitek, A., 2012. Phys. Med. Biol. 57, 6779–6795.
Soares, E.J., Byrne, C.L., Glick, S.J., 2000. IEEE Trans. Med. Imaging 19, 261–270.
Soares, E.J., Glick, S.J., Hoppin, J.W., 2005. IEEE Trans. Med. Imaging 24, 112–121.
Stayman, J.W., Fessler, J.A., 2004. IEEE Trans. Med. Imaging 23, 1543–1556.
Teune, L.K., Bartels, A.L., de Jong, B.M., Willemsen, A.T.M., Eshuis, S.A., de Vries, J.J., van Oostrom, J.C.H., Leenders, K.L., 2010. Mov. Disord. 25, 2395–2404.
the Alzheimer's Disease Neuroimaging Initiative, Liu, M., Cheng, D., Wang, K., Wang, Y., 2018. Neuroinformatics 16, 295–308.
Tripathi, Madhavi, Tripathi, Manjari, Damle, N., Kushwaha, S., Jaimini, A., D'Souza, M.M., Sharma, R., Saw, S., Mondal, A., 2014. Neuroradiol. J. 27, 13–21.
Wilson, D.W., Tsui, B.M.W., Barrett, H.H., 1994. Phys. Med. Biol. 39, 847–871.
N.d. 87.