Individual Analysis of Molecular Brain Imaging Data Through Automatic Identification of Abnormality Patterns

Ninon Burgos, Jorge Samper-González, Anne Bertrand, Marie-Odile Habert, Sébastien Ourselin, Stanley Durrleman, M. Jorge Cardoso, Olivier Colliot

HAL Id: hal-01567343
https://hal.inria.fr/hal-01567343

Submitted on 22 Jul 2017

Individual Analysis of Molecular Brain Imaging Data Through Automatic Identification of Abnormality Patterns

Ninon Burgos[1,2], Jorge Samper-González[1,2], Anne Bertrand[1,2,3],
Marie-Odile Habert[4], Sébastien Ourselin[5,6], Stanley Durrleman[1,2],
M. Jorge Cardoso[5,6], and Olivier Colliot[1,2,3,7]

[1] Inria Paris, Aramis project-team, Paris, France
[2] Sorbonne Universités, UPMC Univ Paris 06, Inserm, CNRS, Institut du Cerveau et la Moelle épinière (ICM) - Pitié-Salpêtrière Hospital, Paris, France
[3] AP-HP, Department of Neuroradiology, Pitié-Salpêtrière Hospital, Paris, France
[4] AP-HP, Department of Nuclear Medicine, Pitié-Salpêtrière Hospital, Sorbonne Universités, UPMC Univ Paris 06, Inserm U 1146, CNRS UMR 7371, Laboratoire d'Imagerie Biomédicale, Paris, France
[5] Translational Imaging Group, CMIC, University College London, London, UK
[6] Dementia Research Centre, University College London, London, UK
[7] AP-HP, Department of Neurology, Pitié-Salpêtrière Hospital, Paris, France

Abstract. We introduce a pipeline for the individual analysis of positron emission tomography (PET) data on large cohorts of patients. This pipeline consists for each individual of generating a subject-specific model of healthy PET appearance and comparing the individual's PET image to the model via a novel regularised Z-score. The resulting voxel-wise Z-score map can be interpreted as a subject-specific abnormality map that summarises the pathology's topographical distribution in the brain. We then propose a strategy to validate the abnormality maps on several PET tracers and automatically detect the underlying pathology by using the abnormality maps as features to feed a linear support vector machine (SVM)-based classifier.

We applied the pipeline to a large dataset comprising 298 subjects selected from the ADNI2 database (103 cognitively normal, 105 late MCI and 90 Alzheimer's disease subjects). The high classification accuracy obtained when using the abnormality maps as features demonstrates that the proposed pipeline is able to extract for each individual the signal characteristic of dementia from both FDG and Florbetapir PET data.

1 Introduction

Long before the clinical symptoms of the disease appear, neuroimaging, mainly magnetic resonance (MR) and positron emission tomography (PET), plays an important role in the diagnosis of dementia [1]. Information derived from PET images is of crucial value: ^{18}F-fluorodeoxyglucose (FDG) PET reflects the glucose consumption, which correlates with the activity of the synapses, while other PET

tracers such as Florbetapir are used to image the deposition of beta-amyloid (Aβ) plaques in the brain. However, the analysis of multiple imaging modalities for diagnostic purposes is to date challenging, and hardly translated to clinical practice. The main drawback is represented by the large amount of information that needs to be consistently processed and analysed to derive clinically useful information.

A popular way to extract meaningful information from neurological images is to use computational methods based on machine learning to directly estimate the category of pathology in a patient. Most machine learning methods developed for classification in dementia studies extract the features used to draw the border that differentiates normality from abnormality directly from the images, e.g. thickness of the cortex extracted from structural MR images [2], or glucose consumption extracted from PET images [3]. However, these features are affected by the anatomical variability present in the population, which acts as a confounding factor making the task of finding the frontier (i.e. the decision function) between normality and abnormality very challenging. Instead of trying to find this frontier at the population level, transporting the problem to the individual level might reduce its complexity.

In previous work, we developed a framework for the analysis of FDG PET data that consists of creating a patient-specific model of healthy PET appearance and comparing the patient's PET image to the model via a Z-score, thus providing voxel-wise statistics on the variation of glucose metabolism in a control population [4]. We showed that this approach was able to distinguish subgroups in a small dataset comprising 22 subjects with distinct neurodegenerative syndromes [4].

In this paper, we introduce a pipeline for the individual analysis of PET data on large cohorts of patients. This pipeline consists of generating a subject-specific model of healthy PET appearance for each subject following the method described in [4] and comparing the subject's PET image to the model via a novel regularised Z-score, which results in the generation of subject-specific abnormality maps summarising the pathology's topographical distribution in the brain. We then propose a strategy to validate the abnormality maps on several PET tracers and automatically detect dementia by using the abnormality maps as features to feed a linear support vector machine (SVM)-based classifier. This strategy enables us to assess on a large dataset composed of 298 subjects selected from the ADNI2 database if the proposed subject-specific abnormality maps are able to extract for each individual the signal characteristic of abnormality from both FDG and Florbetapir PET data, with the aim to reduce the confounding impact of anatomical variability when trying to distinguish disease versus normal ageing.

2 Methods

2.1 Data

Imaging data were obtained from the ADNI2 database[1]. We selected 298 participants who had T1-weighted MRI, ^{18}F-FDG PET, and Florbetapir (^{18}F-AV45) PET images at baseline and were diagnosed as cognitively normal (CN) (n=103, 68 Aβ^-), late mild cognitive impairment (LMCI) (n=105, 71 Aβ^+) or Alzheimer's disease (AD) (n=90, 80 Aβ^+). In [5], Landau et al. categorised subjects as amyloid positive (Aβ^+) or negative (Aβ^-) on Florbetapir based on a cortical mean cutoff of 1.11. Here, we define as amyloid positive the subjects with a cortical mean standardised uptake value ratio (SUVR) higher than $1.11 + 5\%$ and as amyloid negative the subjects with a cortical mean SUVR lower than $1.11 - 5\%$. The control dataset used in this paper is composed of the CN participants amyloid negative (n=68).

2.2 Data preprocessing

PET images were downloaded from the ADNI website after pre-processing (frame averaging, spatial alignment, interpolation to a standard voxel size, and smoothing to a common resolution of 8 mm full width at half maximum). For each subject, the T1 image was mapped to the PET images using a rigid transformation. The T1 images from all the subjects were then mapped to a common coordinate frame via an affine groupwise registration [6]. Finally, the transformations were applied to the T1, FDG PET and AV45 PET images by updating their image coordinate system (without resampling), forming a database of T1 and PET images globally aligned in a common space.

2.3 Subject-specific analysis of PET data

The proposed subject-specific PET analysis framework consisted of selecting in the control dataset the subjects that were morphologically the most similar to the subject being analysed, creating subject-specific models of healthy PET uptake from the selected controls and the target subject's T1 image, and using the resulting model to create subject-specific abnormality maps.

Selection based on global and local image similarity measures Subjects were first selected from the control dataset according to their global morphological similarity to the target subject, as assessed by a global similarity measure, the normalised cross-correlation (NCC). Because all the subjects were pre-aligned with each other, the T1 image of each subjects was simply resampled to the common space and the NCC was computed between each resampled control subject and the resampled target subject. The 50 control subjects with the highest NCC

---

[1] Imaging data were provided by the Alzheimer's disease neuroimaging initiative (http://adni.loni.ucla.edu/).

were selected (top 75%). This step is meant to discard the controls too dissimilar to the target and thus limit the computational time while maintaining a high synthesis accuracy.

The T1 images of the 50 pre-selected controls were then non-rigidly registered to the target subject's T1 image in its native space [7], and the PET images of the control dataset, pre-aligned to the T1 images, were mapped using the same transformation to the target subject. Once non-rigidly aligned to the target subject, the controls morphologically the most similar to the target subject at the voxel level were identified using a local image similarity measure, the structural image similarity (SSIM) [8].

Subject-specific models of healthy PET appearance To generate the subject-specific model, which is composed of two elements: a spatially-varying weighted average and a spatially-varying weighted standard deviation, the controls locally selected were fused based on their morphological similarity to the target subject. The weights, corresponding to the contribution of each control subject to the model, were obtained by ranking at each voxel x the SSIM across the N globally pre-selected control subjects and applying an exponential decay function: $w_n(x) = e^{-\beta r_n(x)}$, where $r_n(x)$ denotes the rank of the nth control subject, and $\beta = 0.5$ [4]. For each of the N pre-selected subjects in the control dataset, let the nth mapped PET image be denoted by J_n. The two subject-specific model elements (I_μ, I_σ) are computed as follows:

$$
\begin{aligned}
I_\mu(x) &= \frac{\sum_{n=1}^{N} w_n(x) \cdot J_n(x)}{\sum_{n=1}^{N} w_n(x)} \quad, \\
I_\sigma(x) &= \sqrt{\frac{N_w}{N_w - 1} \frac{\sum_{n=1}^{N} w_n(x) \cdot (J_n(x) - I_\mu(x))^2}{\sum_{n=1}^{N} w_n(x)}}
\end{aligned}
\tag{1}
$$

where N_w is the number of non-zero weights.

Subject-specific abnormality maps To compare the target subject's PET image to the subject-specific model, in [4] a Z-score was computed for each voxel of the image. However, we observed that this leads to the generation of high frequency signals in certain areas due to the standard deviation approaching zero. To avoid this problem, we define a regularised Z-score

$$
\tilde{Z}(x) = \frac{I(x) - I_\mu(x)}{I_\sigma(x) + \alpha * I_{\bar{\sigma}}}
\tag{2}
$$

where $I_{\bar{\sigma}}$ is the standard deviation averaged over all the voxels. We set α equal to 2 as a compromise between the resulting Z-score maps being too smooth and the presence of high frequency signals. The voxel-wise regularised Z-score map can be interpreted as an *abnormality map*, as it statistically evaluates the localised deviation of the subject-specific uptake with respect to the healthy uptake distribution.

2.4 Validation scheme

To assess the ability of the abnormality maps to extract relevant information from PET data on a large dataset and to offer a new strategy for computer-assisted diagnosis, we propose to use the abnormality maps as features to feed a linear SVM classifier.

Non-linear alignment to group space A way to compare the abnormality maps, each generated in the subject's native space, across all the subjects, is to align them with each other. As the T1 images from all the subjects were already mapped to a common coordinate frame via an affine groupwise registration, the T1 images were subsequently non-rigidly registered to the group-space. The same transformations were then applied to the abnormality maps.

Linear SVM classifier We chose a linear SVM to classify the abnormality maps. A linear kernel was calculated using the inner product for each pair of abnormality maps available in the dataset (using all the brain voxels). This kernel was then used as input for the generic SVM[2]. Two nested 10-fold cross-validation procedures were used to train the classifier and to optimise the hyperparameters. The process was repeated ten times and the classification results averaged over the ten repeats. This process guarantees an unbiased evaluation of the classification accuracy.

Classification tasks The experiments consisted of two simple tasks:
1. differentiating cognitively normal subjects from subjects with a disease, i.e. CN vs AD and CN vs LMCI;
2. differentiating between amyloid negative and amyloid positive subjects (β^- vs β^+).

For the first experiment, 219 subjects (68 CN Aβ^-, 71 LMCI Aβ^+ and 80 AD Aβ^+) were considered, while for the second experiment 298 subjects (112 Aβ^- and 186 Aβ^+) were analysed.

Comparison to state-of-the-art To set the results in perspective, the subjects' PET images themselves and state-of-the-art Z-maps were also used as features and fed to the classifier. The state-of-the-art Z-maps were obtained by comparing the subject's PET image in the group space to the mean and standard deviation computed from all the 68 subjects in the control dataset, also in the group space.

3 Results

Abnormality maps were generated for each of the 298 ADNI2 participants selected, for both the FDG and AV45 PET images. Note that for the CN β^-
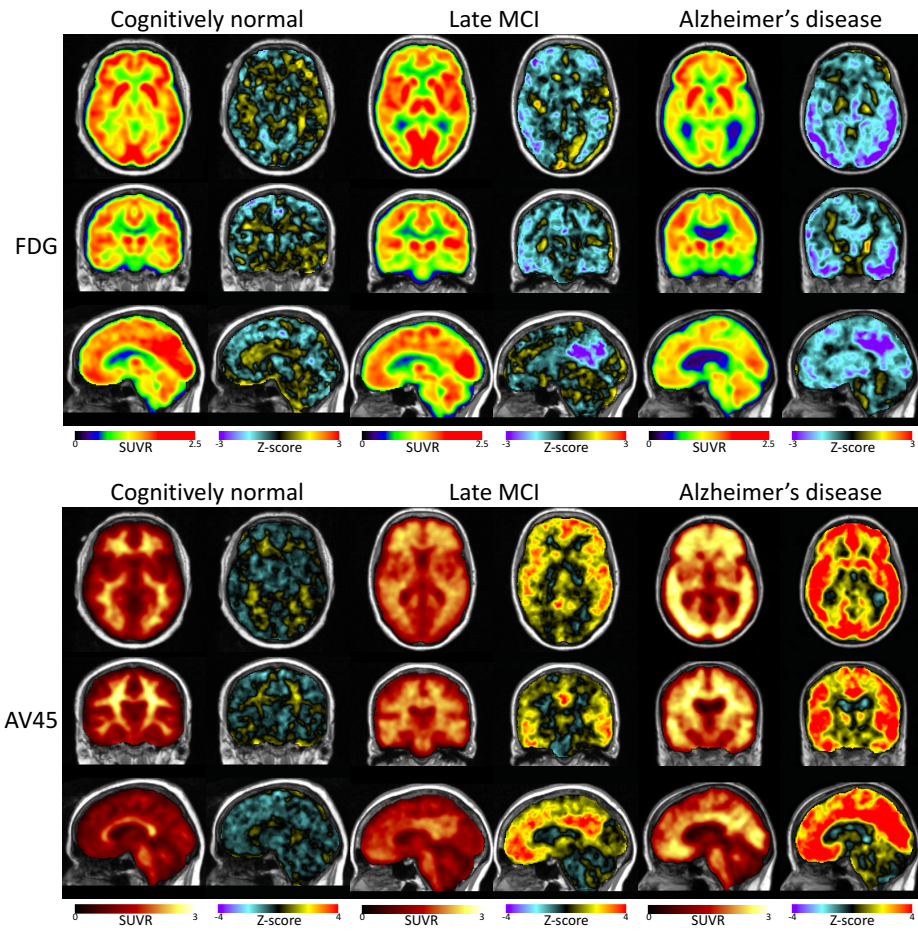
---

[2] http://scikit-learn.org

Fig. 1. Examples of FDG PET images with the corresponding abnormality maps (top) and of AV45 PET images with the corresponding abnormality maps (bottom) for a CN, an LMCI and an AD subject.

subjects (forming the control dataset), a leave-one-out strategy was used, i.e. the images of the CN subject being processed were excluded from the control database. Examples of abnormality maps are displayed in Figure 1 for a CN, a late MCI and an AD subject. We observe that, as expected, no specific signal is being detected for the CN subject, for both the FDG and AV45 tracers. For the LMCI subject, abnormal glucose uptake is detected mainly in the precuneus and in the cingulate gyrus, and abnormal amyloid deposition is detected in the frontal, parietal, temporal and cingulate cortices, which is consistent with previous observations [9]. Finally, for the AD subject, abnormal glucose uptake is

Table 1. Balanced accuracy obtained when using PET images, state-of-the-art Z-maps, and the proposed subject-specific abnormality maps as features of the linear SVM classification algorithm. The average ± SD balanced accuracy, obtained over ten repeats, is expressed in percentages.

	FDG			AV45		
	PET	Zmap	Abn. map	PET	Zmap	Abn. map
CN vs AD	88.9 ± 1.1	89.6 ± 1.2	91.6 ± 1.2	100 ± 0	100 ± 0	100 ± 0
CN vs LMCI	78.3 ± 1.7	78.7 ± 1.9	80.5 ± 1.6	100 ± 0	100 ± 0	99.5 ± 0.5
$A\beta^+$ vs $A\beta^-$	71.5 ± 1.0	71.4 ± 1.1	73.9 ± 1.7	99.8 ± 0.4	99.4 ± 0.5	99.7 ± 0.5

detected in areas such as the hypocampus, the precuneus, the cingulate gyrus or the occipital cortex, and abnormal amyloid deposition is detected in all the cortex, which is typical of AD [1].

The abnormality maps were then fed to the linear SVM classifier. The balanced accuracy obtained with the proposed method applied to the FDG data when differentiating CN from AD and LMCI (92% and 81%, respectively) is higher than the balanced accuracy obtained using PET SUVR values (89% and 78%) and the state-of-the-art Z-maps (90% and 79%) as features. Similar results were obtained when differentiating amyloid negative and positive subjects. When analysing AV45 data, using the PET images themselves, the state-of-the-art Z-maps or the proposed abnormality maps leads to similar, highly accurate, classification results. These highly accurate results were expected, but are here confirmed, as differentiating CN ($A\beta^-$) from AD and LMCI (both $A\beta^+$) subjects, or amyloid negative from amyloid positive subjects, based on features extracted from AV45 data is a quite trivial task. More detailed results are shown in Table 1. The high classification accuracy obtained with the abnormality maps confirms their ability to detect meaningful signal from both FDG and AV45 PET images.

To further analyse the classification results, we studied the feature maps generated by the linear SVM classifier that show which voxels are relevant for each classification task.The maps obtained for the classification of CN vs AD subjects and CN vs LMCI subjects are shown in Figure 2. We observe that the areas that were detected as abnormal with the proposed method (i.e hypocampus, precuneus, cingulate gyrus) are also the ones mostly used to separate AD from CN subjects, no matter the features used. We also observe that these areas are more strongly targeted when the abnormality maps are used as features compared to the PET images themselves or the state-of-the-art Zmaps. This confirms the ability of the abnormality maps to detect areas that are characteristic of AD. A similar trend was observed for the classification of CN vs LMCI subjects and for the classification of $A\beta^+$ vs $A\beta^-$ subjects (not shown).
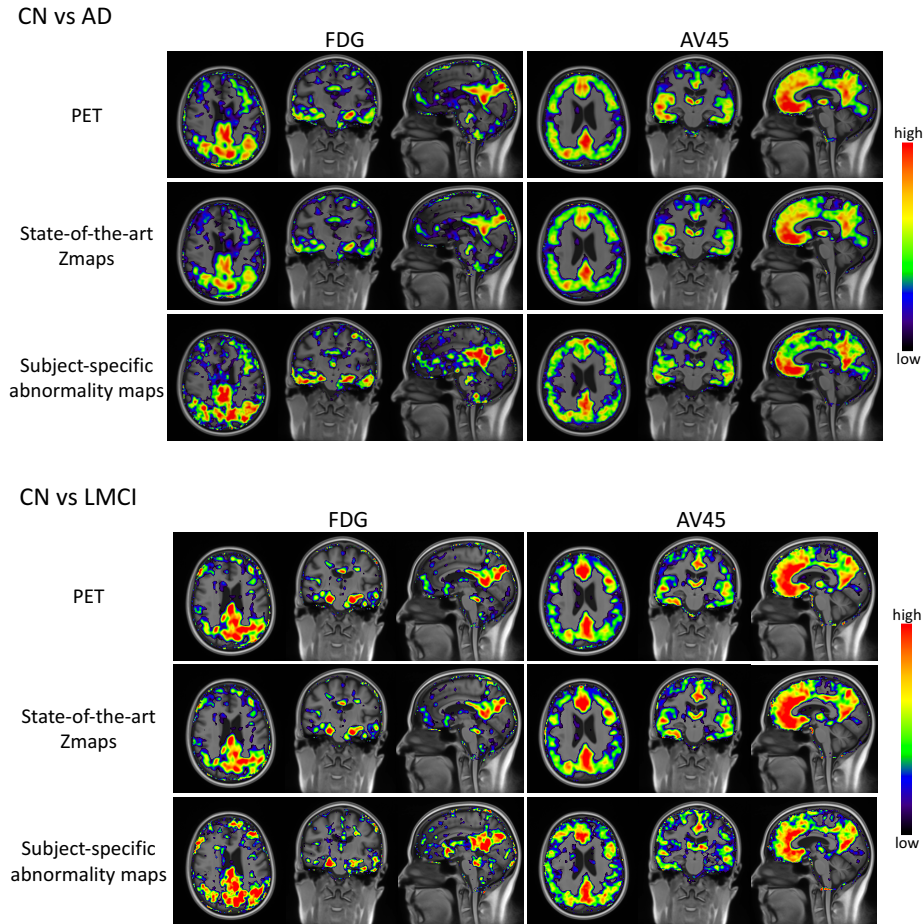
Fig. 2. Voxels the most relevant for the classification of CN vs AD (top) and CN vs LMCI (bottom) subjects when using the PET images themselves, the state-of-the-art Zmaps and the proposed subject-specific abnormality maps as features, for both the FDG and AV45 tracers. The red areas indicate the regions that are mostly used to separate AD or LMCI from CN subjects.

4 Discussion and conclusion

This paper presents a pipeline for the individual analysis of molecular brain images on large-scale datasets. This pipeline is able to automatically locate and characterise the areas characteristic of dementia by generating abnormality maps summarising the pathology's topographical distribution in the brain. This ability was demonstrated by using the abnormality maps as inputs of a classifier and comparing the classification results to the ones obtained when using the

PET images themselves or state-of-the art Z-maps as features. Although the three methods produced a high classification accuracy when differentiating CN from late MCI and AD, and when differentiating amyloid negative subjects from amyloid positive subjects, more accurate results were obtained with the proposed method. These results can be explained by the fact that both the PET images and the state-of-the-art Z-maps are affected by the anatomical variability present in the population, which acts as a confounding factor when trying to differentiate between normality and abnormality. As the subject-specific mean and standard deviation used to compute the abnormality maps are obtained by selecting the subjects in the control dataset that are morphologically the most similar to target subject, the abnormality maps are less impacted by morphological variability.

The abnormality maps have two complementary uses. They can i) help clinicians in their diagnosis by highlighting, in a data-driven fashion, the pathological areas obtained from the individual PET data, and ii) provide quantitative, voxel-based, abnormality scores that can be used as input for computer-assisted diagnosis tools for the automatic detection of dementia.

References

1. Jagust, W.: Positron emission tomography and magnetic resonance imaging in the diagnosis and prediction of dementia. Alzheimer's & Dementia **2**(1) (2006) 36–42
2. Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., Fox, N.C., Jack, C.R., Ashburner, J., Frackowiak, R.S.: Automatic classification of MR scans in Alzheimer's disease. Brain **131**(3) (2008) 681–689
3. Illán, I., Górriz, J., Ramírez, J., Salas-Gonzalez, D., López, M., Segovia, F., Chaves, R., Gómez-Rio, M., Puntonet, C.G., the Alzheimer's Disease Neuroimaging Initiative: ^{18}F-FDG PET imaging analysis for computer aided Alzheimer's diagnosis. Information Sciences **181**(4) (2011) 903–916
4. Burgos, N., Cardoso, M.J., Mendelson, A., Schott, J.M., Atkinson, D., Arridge, S.R., Hutton, B.F., Ourselin, S.: Subject-specific models for the analysis of pathological FDG PET data. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. (2015) 651–658
5. Landau, S.M., Mintun, M.A., Joshi, A.D., Koeppe, R.A., Petersen, R.C., Aisen, P.S., Weiner, M.W., Jagust, W.J.: Amyloid deposition, hypometabolism, and longitudinal cognitive decline. Annals of Neurology **72**(4) (2012) 578–586
6. Modat, M., Cash, D.M., Daga, P., Winston, G.P., Duncan, J.S., Ourselin, S.: A symmetric block-matching framework for global registration. Proc. SPIE, Medical Imaging **9034** (2014)

7. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. Computer methods and programs in biomedicine **98**(3) (2010) 278–84

8. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4) (2004) 600–612

9. Forsberg, A., Engler, H., Almkvist, O., Blomquist, G., Hagman, G., Wall, A., Ringheim, A., Långström, B., Nordberg, A.: PET imaging of amyloid deposition in patients with mild cognitive impairment. Neurobiology of Aging **29**(10) (2008) 1456–1465