New advances in the Clinica software platform for clinical neuroimaging studies

Alexandre Routier, Arnaud Marcoux, Mauricio Diaz Melo, Jérémy Guillon, Jorge Samper-González, Junhao Wen, Simona Bottani, Alexis Guyot, Elina Thibeau-Sutre, Marc Teichmann, et al.

New advances in the Clinica software platform for clinical neuroimaging studies

Alexandre Routier[1], Arnaud Marcoux[1], Mauricio Diaz Melo[2], Jérémy Guillon[1], Jorge Samper-González[1], Junhao Wen[1], Simona Bottani[1], Alexis Guyot[1], Elina Thibeau–Sutre[1], Marc Teichmann[3,4], Marie-Odile Habert[5,6,7], Stanley Durrleman[1], Ninon Burgos[1], Olivier Colliot[1,8]

[1] ARAMIS Lab, ICM, Inserm U1127, CNRS UMR 7225, Sorbonne University, Inria, Paris, France
[2] Inria Paris, SED, 75013 Paris, France
[3] FrontLab, ICM, Inserm U1127, CNRS UMR 7225, Sorbonne University, Paris, France
[4] Department of Neurology, National Reference Center for 'PPA and rare dementias', Institute for Memory and Alzheimer's Disease, Hôpital de la Pitié-Salptrière Hospital, AP-HP, Paris. France
[5] Laboratoire d'Imagerie Biomédicale, Sorbonne Université, Inserm U 1146, CNRS UMR 7371, Paris, France
[6] AP-HP, Hôpital de la Pitié-Salpêtrière Hospital, Department of Nuclear Medicine, Paris, France
[7] Centre Acquisition et Traitement des Images (CATI, cati-neuroimaging.com)
[8] AP-HP, Departments of Neuroradiology and Neurology, Hôpital de la Pitié-Salpêtrière Hospital, Paris, France

Introduction

Carrying out clinical neuroscience studies usually involves many data analysis steps including image preprocessing, extraction of image-derived measurements and statistical analysis. Most of these steps also need to be performed when developing machine learning methods for neuroimaging. Clinica aims to automate the processing and statistical analysis of neuroimaging data and ease the development of machine learning approaches. We present the new functionalities and improvements made to Clinica, including framework upgrades, continuous integration and new pipelines (structural connectome, surface-based PET analysis, fMRI preprocessing, machine learning).

Methods

Clinica is based on the Nipype library (Gorgolewski et al., 2011) and relies on tools from the neuroimaging community to build pipelines (FreeSurfer (Fischl, 2012), FSL (Jenkinson et al., 2012), SPM (Frackowiak et al., 1997), ANTs (Avants et al., 2014), MRtrix3 (Tournier et al., 2012), and PETPVC (Thomas et al., 2016)) or build machine learning modules (scikit-learn (Pedregosa et al., 2011)). The only assumption on the raw data is that they follow the BIDS format (Gorgolewski et al., 2016). New functionalities of Clinica can be divided into three main parts.

First, new pipelines were developed in this release of Clinica:

- `fmri-preprocessing`: This pipeline corrects fMRI datasets for head motion, slice-timing, susceptibility distortion and mainly relies on SPM. It also performs coregistration with T1, spatial normalization into MNI space and spatial smoothing.

- `dwi-connectome`: This pipeline computes a structural connectome, i.e. a weighted graph encoding anatomical connections between a set of brain regions. It relies on the constrained spherical deconvolution diffusion model implemented in MRtrix (Tournier et al., 2012), performs probabilistic tractography and computes a connectome using the Desikan & Destrieux atlases from FreeSurfer.

- `pet-surface` (Marcoux et al., 2018): This pipeline enables the analysis of PET data on the cortical surface. It includes co-registration of PET and T1 MR images, intensity normalization, partial volume correction, robust projection of the PET signal onto the cortical surface extracted from the coregistered T1 image, and spatial normalization to a template.

- `machinelearning-prepare-spatial-svm` (Cuingnet et al., 2013): This pipeline prepares T1 MRI and PET data for performing classification with an SVM that accounts for the spatial and anatomical structure of neuroimaging data, leading to more regular and anatomically interpretable classification results.

The pipelines providing image-derived measurements are listed in **Fig 1**. Secondly, the core of Clinica was consolidated and upgraded to use Python 3 and Nipype 1. Thirdly, we developed a robust continuous integration framework, addressing some specific challenges raised by medical image processing. For instance, pipelines including stochastic models make it difficult to reproduce any experiment with machine precision. To solve this problem, we implemented comparison criteria adapted for each type of data and processing pipelines (correlation coefficient, metric to control the quantity of voxels exceeding certain relative difference thresholds). This allows developers to monitor the consistency of their results with respect to their developments or after a software package upgrade.

Results

The new pipelines expand the scope of modalities covered by Clinica and offer new functionalities. **Fig 2** illustrates how pipelines can be chained to carry out a study, from the conversion of the raw imaging data to the statistical analysis.

Clinica has already been used for clinical research studies (Routier et al., 2018; Wen et al., 2018) and to develop machine learning frameworks for Alzheimer's disease diagnosis (Samper-González et al., 2018; Toro et al., 2018). This is in line with the target audience of Clinica, namely neuroscientists and clinicians conducting neuroimaging studies, and researchers in machine learning.
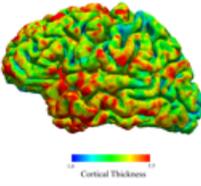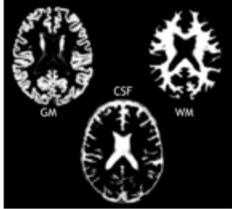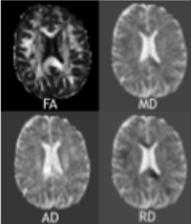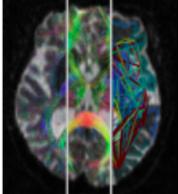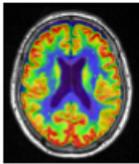
Conclusions

Clinica (www.clinica.run) is an open source software platform for reproducible clinical neuroscience studies. It aims to make clinical research studies easier and pursues the community effort of reproducibility.

References

Avants, B. B.., et al. (2014). The Insight ToolKit image registration framework. Front Neuroinform 8.

Cuingnet, R. (2013). Spatial and Anatomical Regularization of SVM: A General Framework for Neuroimaging Data. IEEE Trans Pattern Anal Mach Intell 35, 682696.

Fischl, B. (2012). FreeSurfer. NeuroImage 62, 774-781.

Gorgolewski, K. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. Front. Neuroinform. 5.

Gorgolewski, K. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Scientific Data 3, 160044.

Jenkinson, M., et al. (2012). FSL. Neuroimage 62, 782-790. doi:10.1016/j.neuroimage.2011.09.015.

Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825-2830.

Routier, A. (2018). Structural, Microstructural, and Metabolic Alterations in Primary Progressive Aphasia Variants. Frontiers in Neurology 9.

Samper-González, J. (2018). Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. NeuroImage.

Toro, C. (2018). Supervoxels-Based Histon as a New Alzheimer's Disease Imaging Biomarker. Sensors 18, 1752.

Thomas, B. A., et al. (2016). PETPVC: a toolbox for performing partial volume correction techniques in positron emission tomography. Physics in Medicine and Biology 61, 79757993.

Tournier, J.-D., et al. (2012). MRtrix: Diffusion tractography in crossing fiber regions. Int. J. Imaging Syst. Technol. 22, 5366.

Wen, J., Zhang, H., Alexander, D.C., Durrleman, S., Routier, A., Rinaldi, D., Houot, M., Couratier, P., Hannequin, D., Pasquier, F., Zhang, J., Colliot, O., Le Ber, I., Bertrand, A., 2018. Neurite density is reduced in the presymptomatic phase of C9orf72 disease. JNNP.

Figures

A N A T O M I C A L M R I	**t1-freesurfer-∗** Cortical surface, cortical thickness, subcortical structures Dependencies: FreeSurfer • Surface-based features (Cortical Thickness) • Regional features (Cortical Thickness) using atlases (Desikan, Destrieux)	**t1-volume-∗** Tissue segmentation (GM, WM, CSF), normalization to standard space (MNI) Dependencies: SPM, CAT12 • Voxel-based features (GM, WM, CSF) • Regional features (GM) using atlases (AAL2, AICHA, Hammers, LPBA40, Neuromorphometrics)
D I F F U S I O N M R I	**dwi-dti** Diffusion tensor imaging (DTI), extraction of FA, MD, AD & RD and normalization to standard space (MNI) Dependencies: FSL, ANTs, MRtrix3 • Voxel-based features (FA, MD, AD, RD) • Regional features (FA, MD, AD, RD) using atlases (JHUDTI81, JHUTracts)	**dwi-connectome** Tractography & connectome Dependencies: FreeSurfer, FSL, MRtrix • Probabilistic tractography • Structural connectome using atlases (Desikan, Destrieux)
P E T	**pet-volume** Registration to T1 MRI, partial volume correction, intensity normalization, spatial normalization to standard space (MNI) Dependencies: SPM, PETPVC, CAT12 • Voxel-based features (FDG, Amyloid) • Regional features (FDG, Amyloid) using atlases (AAL2, AICHA, Hammers, LPBA40, Neuromorphometrics)	**pet-surface** Projection of PET acquisition into surface of the cortex Dependencies: FreeSurfer, FSL, SPM, PETPVC • Surface-based features (FDG, Amyloid) • Regional features (FDG, Amyloid) using atlases (Desikan, Destrieux)
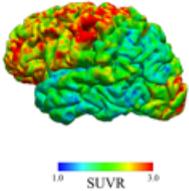
Figure 1 List of pipelines providing image-derived measurements available in Clinica with their dependencies and outputs.
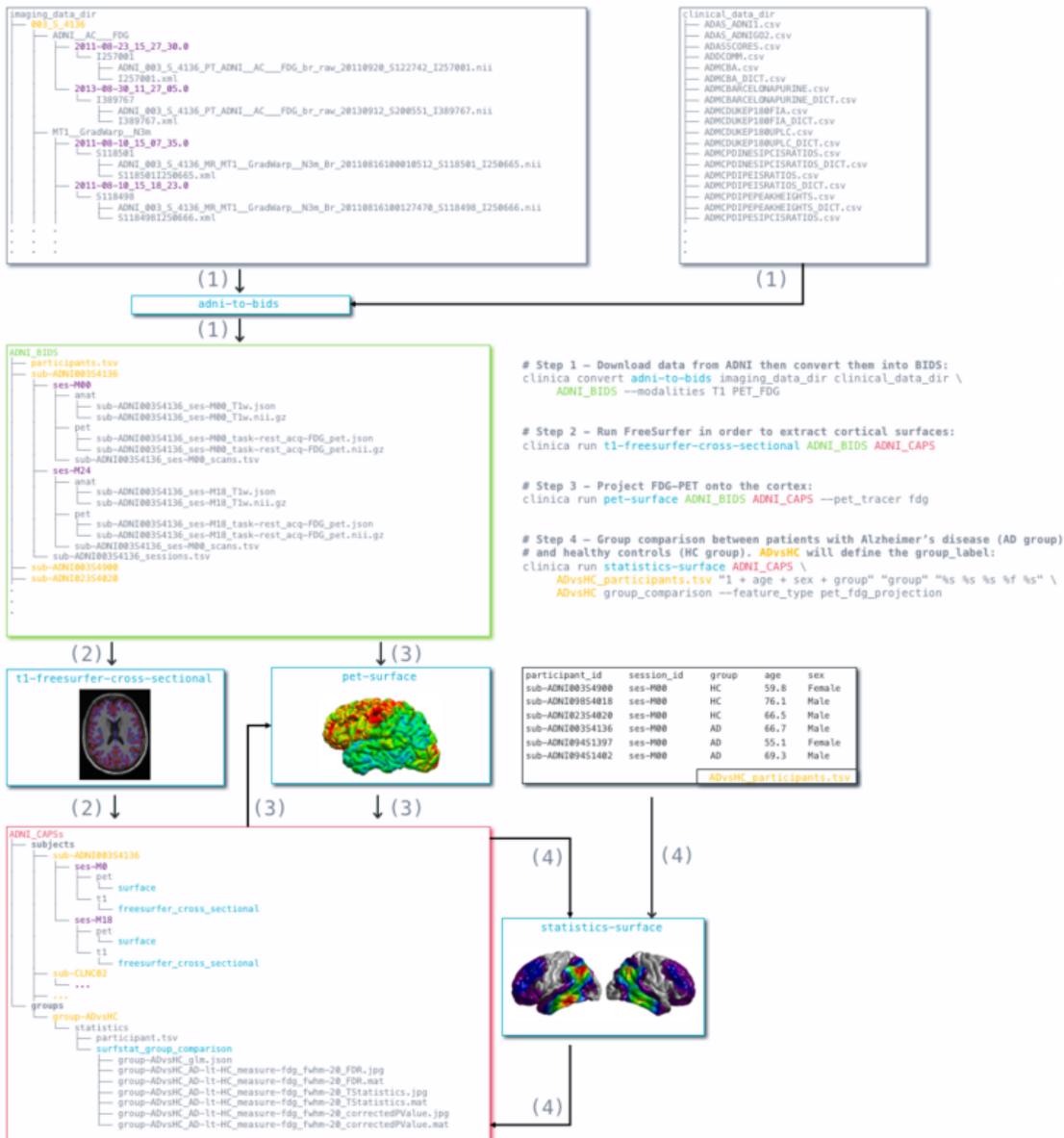
Figure 2 Diagram illustrating the Clinica pipelines involved when performing a group comparison of FDG PET data projected on the cortical surface between patients with Alzheimer's disease and healthy controls from the ADNI database. First, clinical and neuroimaging data are downloaded from the ADNI website and data are converted into BIDS with the adni-to-bids tool from Clinica (1). Estimation of the cortical and white surface is then produced by the t1-freesurfer-cross-sectional pipeline in a single command line (2). Afterwards, FDG PET data can be projected on the subject's cortical surface and normalized to the FsAverage template from FreeSurfer using the pet-surface pipeline (3). Finally, a TSV file with demographic information of the population studied is given to the statistics-surface pipeline to generate the results of the group comparison (4).