Pipeline Comparison for the Pre-Processing of Resting-State Data in Epilepsy

Bianca De Blasi
Dept. of Medical Physics
University College London
London, UK
bianca.blasi.15@ucl.ac.uk

Matthias Koepp
Inst. of Neurology
University College London
London, UK
m.koepp@ucl.ac.uk

Ilaria Boscolo Galazzo
Dept. of Computer Science
University of Verona
Verona, Italy
ilaria.boscologalazzo@univr.it

Anna Barnes
Inst. of Nuclear Medicine
University College London
Hospitals
London,UK
anna.barnes1@nhs.net

Luca Pasetto
Dept. of Computer Science
University of Verona
Verona, Italy
luca.pasetto_01@studenti.univr.it

Gloria Menegaz
Dept. of Computer Science
University of Verona
Verona, Italy
gloria.menegaz@univr.it

Silvia Francesca Storti
Dept. of Computer Science
University of Verona
Verona, Italy
silviafrancesca.storti@univr.it

Abstract—**Noise removal is a critical step to recover the signal of interest from resting-state fMRI data. Several pre-processing pipelines have been developed mainly based on nuisance regression or independent component analysis. The aim of this work was to evaluate the ability in removing spurious non-BOLD signals of different cleaning pipelines when applied to a dataset of healthy controls and temporal lobe epilepsy patients. Increased tSNR and power spectral density in the resting-state frequency range (0.01-0.1 Hz) were found for all pre-processing pipelines with respect to the minimally pre-processed data, suggesting a positive gain in terms of temporal properties when optimal cleaning procedures are applied to the acquired fMRI data. All the pre-processing pipelines considered were able to recover the DMN through group ICA. By visually comparing this network across all the pipelines and groups, we found that AROMA, SPM12, FIX and FIXMC were able to better delineate the posterior cingulate cortex.**

Keywords—resting-state fMRI, pre-processing pipelines, epilepsy, default mode network

I. INTRODUCTION

Resting-state functional magnetic resonance imaging (rs-fMRI) based on the blood-oxygenated-level dependent (BOLD) contrast is a powerful tool for studying spontaneous brain activity in healthy controls and patients [1], [2]. This technique is relatively easy to acquire, it does not depend on a task to be performed and it has been shown to be stable across subjects [3]. However, by acquiring the data in the absence of any task (i.e. at rest), the a-priori information of brain activation is missing. Therefore, these data are not used to localise a brain area active during a specific task (as in task-based fMRI), but rather to investigate brain connectivity from a broader point of view, using functional connectivity (FC) analyses, such as independent component analysis (ICA) or seed-based correlation.

Despite the recent technical improvements in the acquisition phase, the BOLD signal is generally quite noisy. Therefore, it is necessary to recover and separate the signal of interest (related to brain function) from the other noise-related fluctuations in order to obtain more reliable FC results [2], [4]. To achieve artefacts removal, several pre-processing pipelines have been developed which are generally based on nuisance regression or on ICA [5]. These pipelines result in cleaned fMRI time series that more accurately reflect the underlying brain fluctuations of interest and allow to reduce possible bias in the connectivity estimates due to noise confounds. In nuisance regression-based pipelines, motion parameters are estimated during realignment and are used, together with white matter (WM) and cerebral spinal fluid (CSF) average time series, as regressors of no interest [5]. The second group of pipelines exploit ICA, a data-driven method to decompose the fMRI data into signal of interest and structured noise. The actual classification of these components into signal of interest or noise is usually carried out manually, resulting in a time consuming and user-specific procedure. In summary, nuisance-regression techniques regress out variance of no interest in a supervised fashion, while ICA based pipelines do not use data-dependent prior information to determine components of no interest. Recently, different authors have started to propose specific toolboxes for automatically classifying the ICA components, as ICA-AROMA [6] or FIX [7]. However, regardless the nature of the denoising pipelines adopted, their effectiveness needs to be assessed in both the control and patient populations.

The aim of this paper was to evaluate the ability of removing spurious non-BOLD signals of different cleaning pipelines, keeping in mind that understanding which share of the resting-state signal is related to neural signals and which proportion is unambiguously related to noise is a challenging question [8]. These were applied to a dataset of healthy controls and temporal lobe epilepsy patients in order to assess their behaviour in physiological and pathological conditions. The resulting cleaned data were compared by computing temporal SNR (tSNR) and power spectra density measures. Additionally, we investigated the accuracy of each cleaned dataset to recover the Default Mode Network (DMN) through

group-based ICA, aiming at providing preliminary findings about the impact of the cleaning procedure on network-based connectivity.

II. METHODS

A. Population

Nineteen temporal lobe epilepsy patients (nine left-sided [LTLE], 41 ± 14y) and twenty age and gender matched healthy controls (38 ± 10y) were enrolled. This study was approved by the London - Stanmore Research Ethics Committee (REC reference 15/LO/1051). Written informed consent was obtained for each subject.

B. Image Acquisition

Imaging was carried out on a 3T Siemens PET/MRI scanner equipped with a 16-channel head and neck coil. During the rs-fMRI acquisition, the subjects were instructed to stay still and relaxed and to close their eyes without falling asleep. rs-fMRI data were acquired using an echo-planar imaging sequence and the following parameters: TR=2020 ms, TE=30 ms, flip angle=70°, voxel size=3x3x4 mm^3, 36 slices, 260 volumes. High resolution 3D T1-weighted anatomical images were also acquired: TR=2000 ms, TE=2.92 ms, voxel size=1.1x1.1x1.1 mm^3, dimensions=193x187x173.

C. Raw Data and Basic Pre-processing

Three pipelines were considered to minimally pre-process the data. The first included only the spatial normalization of the raw data to the MNI 2-mm standard space (referred to as Raw). The second (Min) comprised minimal pre-processing steps using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/): realignment, slice timing correction, co-registration to structural, spatial normalisation to MNI 2-mm standard space and spatial smoothing with a 6-mm FWHM kernel. The third (Min+HP) was carried out as the second with the addition of a final high-pass filtering step, with cut-off 0.01 Hz.

D. ICA-AROMA

The fourth pre-processing approach was an ICA-based pipeline in FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) using the ICA-AROMA toolbox to clean each participant's scans specifically from motion confounds (referred to as ICA-AROMA). FSL FEAT pre-processing was initially applied to the data (without temporal filtering) along with MELODIC (automatic estimation of the optimal number of components), as specified in the reference manual. The resulting independent components were then used in ICA-AROMA to automatically identify and remove motion artefacts (non-aggressive option). The cleaned data were finally filtered (high-pass, cut-off 0.01 Hz) and spatially normalized to MNI 2-mm standard space.

E. FIX

Two additional ICA-based approaches (pipelines five and six) were implemented in FSL using the FIX toolbox to clean each participant's scans from various and heterogeneous types of structured noise. Standard FSL FEAT pre-processing was initially applied to the data. Each pre-processed dataset was then decomposed using MELODIC (automatic dimensionality estimation). As FIX requires reliable 'training data' which have to match the data under investigation, manual classification of the components was done on 21 subjects (7 for each of the three groups) accordingly to their spatial distribution, the temporal power spectrum and the time series. This allowed training the classifier which was then applied to all subjects' data. Components automatically classified as artefacts were removed from the data using the non-aggressive option. FIX was applied without (FIX) and with motion regression (FIXMC). In the latter case, the full variance of 24 motion parameters was regressed out. All cleaned data were finally spatially normalised to MNI 2-mm standard space.

F. SPM12

Another pipeline used to pre-process the rs-fMRI data was implemented in SPM12 (referred to as SPM12) and exploited nuisance regression through a general linear model (GLM). The data was pre-processed using realignment, slice timing correction, co-registration to structural data, spatial normalisation to MNI 2-mm standard space. The functional data were finally smoothed by 6-mm FWHM kernel. The normalized T1-weighted scans were segmented and the resulting WM and CSF probability maps were thresholded at 0.9, to strictly retain only pure WM and CSF voxels, and binarised. These masks were then used to extract the WM and CSF average time series to be used in the nuisance regression. WM and CSF signals, together with the 6 motion parameters (estimated in the realignment step), were regressed out from the data. The residuals were finally band-pass filtered (0.0078-0.1 Hz).

G. Temporal Signal-to-Noise Ratio

tSNR can be used to determine the SNR of fMRI time series, by taking into account the mean signal over time. For each pipeline and corresponding cleaned data, tSNR was computed voxel-wise for each subject by dividing the mean signal over time by the standard deviation over time. For each subject, the mean tSNR value in the GM and WM was then computed, using the tissue segmentations previously estimated (probability values >=0.9). The mean was chosen as a summary value to be used in subsequent statistical comparisons.

H. Power Spectral Density

Power spectra were also considered in order to evaluate the impact of the cleaning pipelines on the frequency content, considering the typical frequency range for resting-state (0.01-0.1 Hz). The average time series across the DMN nodes were computed for each subject, relying on the network mask derived on the basis of the ICA template from Smith et al. [10] (the same mask was used for every subject). The DMN was chosen as it represents a robust and reproducible network, encompassing brain regions fluctuating at similar resting-state frequencies, as opposed to the whole GM mask which covers multiple brain areas and thus might be characterized by a

different frequency content. The extracted time series were standardised by subtracting its mean and dividing by its standard deviation. The power spectra were then compute and averaged across subjects. The integral under the curve was also computed between 0.01-0.1 Hz for each pipeline and subject, in order to quantify the power across the range of interest.

I. Statistical Analysis

For each group separately, a one-way analysis of variance (ANOVA) for repeated measures was performed on the tSNR values and on the integral under the curve to test for significant differences across the different pipelines. A post-hoc paired sample two-tailed t-test was applied (p-value<0.05), which was corrected for multiple comparison using Bonferroni correction.

J. ICA Spatial Maps Analysis

Finally, we tested the efficiency of the different cleaning pipelines in terms of ICA networks and spatial maps. A group-based ICA (MELODIC, FSL) was run for every group and pipeline, setting the number of components to 30 (a value commonly used in the literature [9]) to make sure the same number of component was extracted from each group, and the DMN network was again retained for further analyses. The spatial correlation between the group DMN extracted for each pipeline and the DMN template from Smith et al. [10] was assessed using the FSL cross correlation function. Under the assumption that the more the group ICA maps correlate with the corresponding template, the more the cleaning pipeline correctly identifies the true signal [10], we aimed to obtain an initial comparison between the different pre-processing pipelines, in terms of functional connectivity analysis.

III. RESULTS

A. Temporal Signal-to-Noise Ratio

Fig.1 reports the distribution of tSNR values over GM and WM, considering each pipeline and group separately. In all cases, increased tSNR was found for AROMA, FIX, FIXMC and SPM12 pipelines when compared to the minimally pre-processed data. Similar tSNR values were found for SPM12 vs FIXMC, AROMA vs FIX. Additionally, an increased standard deviation was found when the data were processed, as compared to Raw data. Similar tSNR values were found for controls and RTLE, while consistently lower values were shown in the LTLE group.
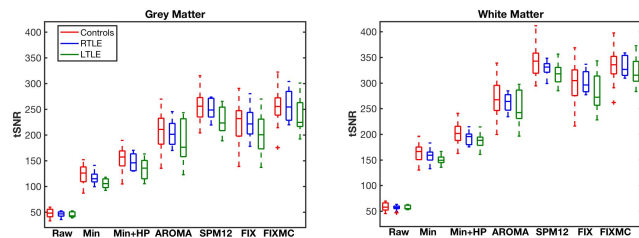

Fig. 1. Distribution of tSNR values in grey matter and white matter, considering each group and pipeline separately.

When statistically compared, for all groups the tSNR values were significantly different across the seven pipelines in both GM and WM (GM: $F_{(6,133)}=164.17$, p-value<0.05 for controls; $F_{(6,63)}=122.68$, p-value<0.05 for RTLE; $F_{(6,56)}=50.26$, p-value<0.05 for LTLE; WM: $F_{(6,133)}=293.85$, p-value<0.05 for controls; $F_{(6,63)}=384.8$, p-value<0.05 for RTLE; $F_{(6,56)}=136.07$, p-value<0.05 for LTLE). The t-test between pairs of pipelines revealed significant differences (p<0.05, Bonferroni corrected) in the tSNR values calculated in GM and in WM for all pairs of pipelines and groups, except in the following cases: AROMA vs FIX for LTLE (GM); SPM12 vs FIX for LTLE and RTLE (GM, WM); and for SPM12 vs FIXMC in all cases (Fig.2).
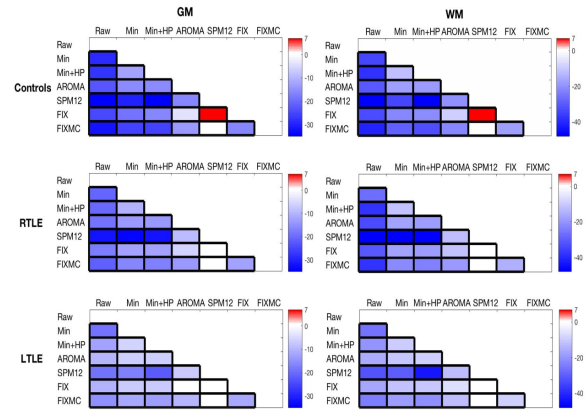

Fig. 2. Graphical representation of the t-values resulting from the statistical comparison of the tSNR measures for each pair of pipelines (reading column first then row). The comparisons which were not statistical significant (p>=0.05) are reported in white.

B. Power Spectral Density

Fig.3 shows the average power spectra of the DMN signals for the three groups and pipelines considered. Of note that the Raw pipeline results are not reported here as they showed the same patterns as the Minimum. In all cases, the Min pipeline presented an initial high peak, possibly related to physiological noise and subject motion, at around 0.002 Hz (out of the range of interest) which was eliminated after processing with the other cleaning pipelines. For each pipeline, the average power spectra followed similar trends across the three groups, with consistent patterns to those usually reported in literature for the DMN network.

Overall, the integral under the curve in the range of frequencies of interest (0.01-0.1 Hz) was always higher for SPM12, AROMA, FIX and FIXMC when compared to the minimally pre-processed data (Raw, Minimum, Minimum+HP). Considering the ANOVA results, for all groups the integral under the curve of the power spectral density was significantly different across all the pipelines ($F_{(6,133)}=51.25$, p-value<0.05 for controls; $F_{(6,63)}=34.68$, p-value<0.05 for RTLE; $F_{(6,56)}=19.12$, p-value<0.05 for LTLE).
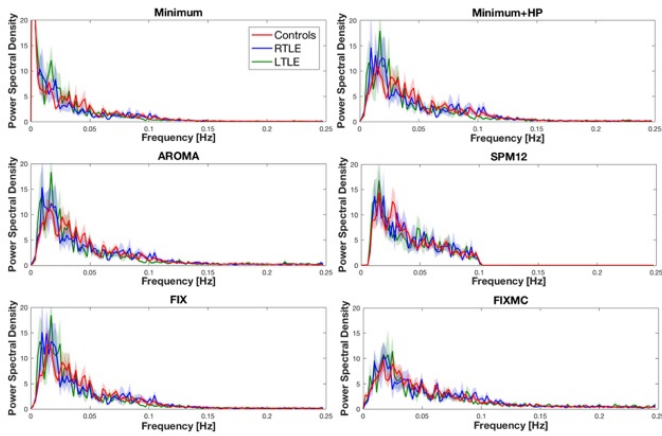
Fig. 3. Average power spectra in the DMN for each group and pipeline. The mean (solid line) is reported together with the standard error of the mean (shading).

The t-test between pairs of pipelines revealed significant differences (p<0.05, Bonferroni corrected) for all pairs of pipelines and groups, except in the following cases: AROMA vs FIX for all groups; AROMA vs FIXMC for RTLE and LTLE; Minimum+HP vs AROMA/FIX/FIXMC for all groups; FIX vs FIXMC for RTLE and LTLE; FIXMC vs Minimum and Raw for the LTLE; Minimum vs Raw for all groups (Fig.4).
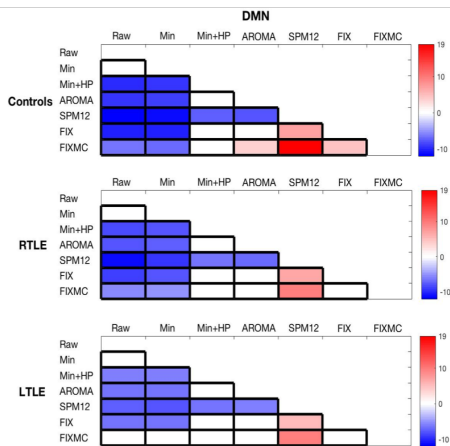

Fig. 4. Graphical representation of the t-values resulting from the statistical comparison of the integral under curve measures in the range of frequencies of interest (0.01-0.1 Hz) (reading column first then row). The comparison which were not statistical significant (p>=0.05) are reported in white.

C. ICA Spatial Maps Analysis

The group-based DMN maps, resulting from the rs-fMRI data cleaned with the different pipelines, are reported in Figg.5-7. The DMN was accurately recovered by all the cleaned datasets. However, in the case of AROMA, SPM12, FIX and FIXMC the posterior cingulate cortex appeared to be more accurate than in the case of Minimum and Mimimum+HP. In the control group, the frontal portion of the DMN was not accurately identified by SPM12 and FIX. Qualitative differences in the frontal area are also present in

the patient groups, which might also be due to alterations in the DMN connectivity due to their pathologies.
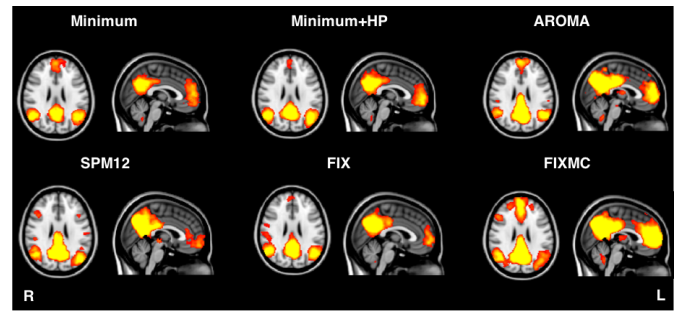

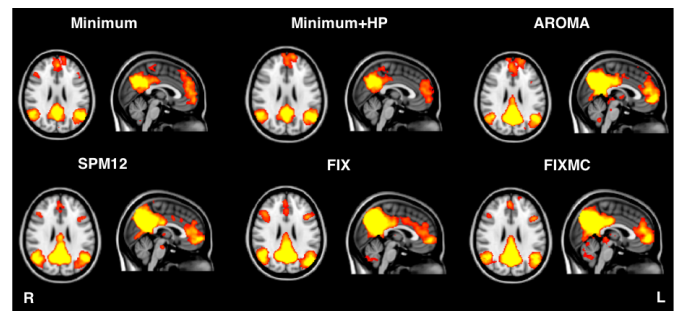Fig. 5. DMN for the control group. The same slice in MNI space has been displayed with a thresholded of z>3.

Fig. 6. DMN for the RTLE group. The same slice in MNI space has been displayed with a thresholded of z>3.

Fig. 7. DMN for the LTLE group. The same slice in MNI space has been displayed with a thresholded of z>3.

Table I reports the values of the spatial correlation between the group DMN obtained from each cleaned dataset and the DMN template. Similar and high correlation values were detected for all the pipelines. A slight increased correlation was found for the Minimum+HP and FIXMC pipelines in the control group; AROMA, SPM12 and FIXMC in RTLE group; FIX in the LTLE group.

TABLE I.   SPATIAL CORRELATION VALUES BETWEEN THE DMN OBTAINED FROM CLEANED DATA AND THE DMN TEMPLATE.

	Controls	RTLE	LTLE
Minimum	0.62	0.57	0.65
Minimum+HP	0.69	0.58	0.63
AROMA	0.65	0.66	0.65
SPM12	0.65	0.66	0.63
FIX	0.64	0.64	0.66
FIXMC	0.67	0.67	0.63

IV. DISCUSSION

The aim of this work was to evaluate the ability of seven different cleaning pipelines in removing spurious non-BOLD signals. These were applied to a dataset of healthy controls and temporal lobe epilepsy patients and the resulting cleaned data were compared, by relying on temporal analyses.

As expected, we found an increased tSNR when the data were pre-processed with SPM12, AROMA, FIX and FIXMC as compared to the minimally pre-processed data. In particular, SPM12 and FIXMC did not show a significant difference, which can be attributed to the fact that both pipelines carry out a regression step where noisy components are completely eliminated (regressed out) from the data. Compared to the raw data, we found an increased standard deviation when any of the other pipelines was used. This was in agreement with previous studies on pipeline comparison and it was attributed to the fact that noise tends to mask the differences within a group of subjects [2], [9]. Therefore, by cleaning the data, one can capture the differences between the subjects in a given group (increased standard deviation), previously covered by the unwanted confounds.

Power spectral density has been used to assess the power in the range of frequency of interest, thereby providing an additional measure of comparison. Previous works mainly qualitatively assessed the power spectral density at high and low frequencies to see how it was changing after appropriate pre-processing [9]. In this work, a quantification of the power spectral density was provided by computing the integral under the curve in the frequency range 0.01-0.1 Hz, of interest in resting-state analysis. This measure was used to compare the different cleaning methods. The Raw and Minimum pipelines showed very similar spectra with an initial peak, out of the frequency range, which was attributed to physiological noise and motion. Indeed, this peak was removed in all the other pipelines which also showed high power content in the range of interest. In terms of power quantification in the frequency range of interest, SPM12 reported the highest power which was significantly different from all the other pipelines. This can be attributed to the regression of WM and CSF, carried out by this pipeline only.

All the pre-processing pipelines considered were able to recover the DMN through group-based ICA.

By visually comparing this network across all pipelines and groups, we found that AROMA, SPM12, FIX and FIXMC were able to better delineate the posterior cingulate cortex. We also reported differences in the anterior areas of the brain in the patients, which might be related to DMN changes due to their pathology. However, when considering the results of the correlation with the DMN template we did not find notable differences across the different pipelines.

V. CONCLUSIONS

To conclude, we reported a comparison of different pre-processing pipelines in terms of tSNR, power spectral density and ability to recover the DMN, for the analysis of rs-fMRI data in healthy controls and epilepsy patients. In this context, it is challenging to define an appropriate ground truth to determine the best pipeline to adopt. On the other hand, it is fundamental to investigate the impact of different pre-processing pipelines on the analysis of resting-state connectivity, especially when considering patient data. In the future, we are planning to extend this work to include more subjects to further confirm these results.

REFERENCES

[1] D. M. Cole, S. M. Smith, C. F. Beckmann, "Advances and pitfalls in the analysis and interpretation of resting-state fMRI data", Frontiers in Systems Neuroscience, vol. 4, pp.1-15, 2010.

[2] O. DiPasquale, et al., "Comparing resting state fMRI de-noising approaches using multi- and single-echo acquisitions", PLoS One, vol. 12, pp.1-25, 2017.

[3] L. Griffanti, et al., "Challenges in the reproducibility of clinical studies with resting state fMRI: an example in early Parkinson's disease", Neuroimage, vol. 154, pp. 704-713, 2016.

[4] C. Caballero-Gaudes, R. C. Reynolds, "Methods for cleaning the BOLD fMRI signal", Neuroimage, vol. 154, pp. 128-149, 2017.

[5] R. H. R. Pruim, M. Mennes, J. K. Buitelaar, C. F. Beckmann, "Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI", Neuroimage, vol. 112, pp. 278-287, 2015.

[6] R. H. R. Pruim, et al., "ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data", Neuroimage, vol. 112, pp. 267–277, 2015.

[7] G. Salimi-Khorshidi, et al., "Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers", Neuroimage, vol. 90, pp. 449-468, 2014.

[8] D. Bzdok, et al., "Formal models of the network co-occurrence underlying mental operations", Plos Comput. Biol., vol. 12, pp.1-31, 2016.

[9] L. Griffanti, et al., "ICA-based artefact removal and accelerated fMRI acquisition for improved resting state imaging", Neuroimage, vol. 95, pp.232-247, 2014.

[10] S. M. Smith, et al., "Correspondence of brain's functional architecture during activation and rest", Proc. Natl. Acad. Sci., vol. 106, pp.13040-13045, 2009.