

# Valid inference for group analysis of functionally aligned fMRI images

## *Inferenza valida per l'analisi di gruppo di immagini fMRI allineate funzionalmente*

Angela Andreella, Riccardo De Santis, and Livio Finos

**Abstract** Functional magnetic resonance imaging (fMRI) data require preprocessing steps before statistical analysis. Multi-subjects fMRI studies are complicated: the brain's anatomical and functional structure varies across subjects. Anatomical alignment does not capture the functional variability across subjects; the functional alignment is then applied. Generally, group analysis on functionally aligned fMRI data refers to between-subject classification. We propose an inference group analysis arguing that using functional aligned images based on Procrustes transformation does not affect type I error.

**Abstract** *I dati di risonanza magnetica funzionale (fMRI) necessitano di fasi di pre-elaborazione prima dell'analisi statistica. Gli studi fMRI multi-soggetto sono complessi: la struttura anatomica e funzionale del cervello varia tra i soggetti. L'allineamento anatomico non cattura la variabilità funzionale tra i soggetti, si applicano dunque metodi di allineamento funzionale. Generalmente, ci si riferisce ad analisi di classificazione quando si parla di studi multi-soggetto su dati allineati funzionalmente. In questo lavoro proponiamo un'analisi di gruppo inferenziale sostenendo che utilizzare immagini allineate funzionalmente tramite metodi basati sulla trasformazione di Procuste non influisce sull'errore di tipo I.*

**Key words:** Procrustes method, ProMises model, fMRI group analysis, fMRI data

---

Angela Andreella  
Department of Economics, University Ca' Foscari Venezia, Italy  
e-mail: [angela.andreella@unive.it](mailto:angela.andreella@unive.it)

Riccardo De Santis  
Department of Statistics, University of Padova, Italy  
e-mail: [riccardo.desantis.1@phd.unipd.it](mailto:riccardo.desantis.1@phd.unipd.it)

Livio Finos  
Department of Developmental Psychology and Socialization, University of Padova, Italy  
e-mail: [livio.finos@unipd.it](mailto:livio.finos@unipd.it)

## 1 Introduction

Functional Magnetic Resonance Image (fMRI) is the most used technique to study the neural underpinnings of human cognition. The brain activation is expressed in the correlation between the sequence of cognitive stimuli and the sequence of measured blood oxygenation levels (BOLD). In reaction to neural activity, changes in brain hemodynamics impact the local intensity of the magnetic resonance signal, i.e., the voxel intensity (single volume elements). The voxel expresses statistically significant neural activity if the correlation computed is statistically significant concerning an inactive region null hypothesis. However, when group analysis is performed, many problems arise.

First, brain activity is not functionally nor anatomically aligned across subjects since brains' anatomical and functional structures vary across subjects, even in response to identical sensory inputs. Anatomical normalization (e.g., [13]) fixes the anatomical misalignment; but it fails to capture the functional variability across subjects. Second, analyzing group activation means performing several statistical tests equal to the number of voxels (i.e., roughly 200,000 statistical tests). It leads to multiple testing problems, and family-wise error rate must be controlled.

In this paper, we proposed a method that simultaneously resolves the functional misalignment and gains power in group analysis inference without affecting the type I error. The functional alignment is done by the ProMises model [2] which can be seen as a procedure that sorts the null hypotheses based on a priori information. The brain images' functional alignment can be described as a sorting criterion independent from the test statistics, making the inference on aligned data completely valid. This procedure is not new in the literature; some methods use, for example, previous findings in similar experiments or the total variance of the variables if central location tests are used [5].

The outline of the paper is as follows. Subsect. 2.1 introduces the ProMises model [2], while Subsect. 2.2 shows the group analysis procedure in fMRI data. The valid inference using functionally aligned data is described in Subsect. 2.3. Finally, the ProMises model is evaluated by performing a group-level inference, analyzing task-related fMRI data in Sect. 3. The entire code used in this work is available in [3].

## 2 Methods

### 2.1 ProMises model

Let  $\{X_i \in \mathbb{R}^{n \times m}\}_{i=1, \dots, N}$  a set of matrices to be aligned and  $M \in \mathbb{R}^{n \times m}$  the shared unknown matrix. The ProMises model uses similarity transformation [7], i.e., scaling, rotation/reflection, and translation, to map  $\{X_i \in \mathbb{R}^{n \times m}\}_{i=1, \dots, N}$  into a common reference matrix  $M$ . The model is defined as follows:

Valid inference for group analysis of functionally aligned fMRI images

$$X_i = \alpha_i(M + E_i)R_i^\top + \mathbf{1}_n^\top t_i \quad (1)$$

where  $E_i$  is the random error matrix with a normal matrix distribution [8]  $E_i \sim \mathcal{M}\mathcal{N}_{nm}(0, \Sigma_n, \Sigma_m)$ ,  $R_i$  is the orthogonal matrix parameter with a von Mises-Fisher distribution [4] (i.e.,  $f(R_i) \sim C(F, k) \exp(kFR_i)$  with  $F \in \mathbb{R}^{m \times m}$  location parameter,  $k \in \mathbb{R}^+$  regularization parameter and  $C(F, k)$  normalizing constant),  $\alpha_i$  is the scaling parameter and  $t_i$  is the translation parameter with  $\mathbf{1}_n$  a  $n$ -dimensional vector of ones. The maximum a posteriori estimates for the sets  $\{R_i\}_{i=1, \dots, N}$  and  $\{\alpha_i\}_{i=1, \dots, N}$  equal

$$\hat{R}_i = \{U_i V_i^\top\}_{i=1, \dots, N}; \quad \hat{\alpha}_{i\hat{R}_i} = \frac{\|\Sigma_m^{-1/2} \hat{R}_i^\top X_i^\top \Sigma_n^{-1/2}\|_F^2}{\text{tr}(D_i)} \quad \forall i \in \{1, \dots, N\}, \quad (2)$$

where  $\|\cdot\|_F$  is the Frobenius norm, and  $U_i D_i V_i^\top$  is the singular value decomposition of  $X_i^\top \Sigma_n^{-1} M \Sigma_m^{-1} + kF$ . For further details about the estimation process, see [2].

## 2.2 Group-level analysis

Let consider the set  $\{X_i \in \mathbb{R}^{n \times m}\}_{i=1, \dots, N}$ . After the  $X_i$  matrices' functional alignment (e.g., using the ProMises model explained in Subsection 2.1), the first-level analysis (i.e., subject-level analysis) is performed to find significant activation in the set of  $m$  voxels for each subject under the null hypothesis of no activation. Therefore, we consider for each subject  $i$  the following model:

$$X_i = DB_i + ZG_i + E_i \quad (3)$$

where  $D \in \mathbb{R}^{n \times p}$  and  $Z \in \mathbb{R}^{n \times q}$  are fixed matrices,  $B_i \in \mathbb{R}^{p \times m}$ ,  $G_i \in \mathbb{R}^{q \times m}$  and  $B_i = B + U_i$ ,  $G_i = G + g_i$  with  $B$  is the true matrix of fixed effects of interest,  $G$  of fixed nuisance effects.  $[U_i^\top | g_i^\top] \sim \mathcal{M}\mathcal{N}(0, \Sigma_m, \Sigma_{pq})$  is the matrix of random effects.

We have now a set of  $N$  matrices  $\{\hat{B}_1, \dots, \hat{B}_N; \hat{B}_i \in \mathbb{R}^{p \times m}\}$ , which for example describes the difference between the neural activation during two stimuli recorded in the voxel  $k \in \{1, \dots, m\}$  of the subject  $i$ . The one-sample t-test [11] is generally performed to study the group's mean activation due to the difference between the neural activation during two stimuli:

$$T = \frac{\hat{\mu}}{\hat{\sigma}/\sqrt{N}}, \quad (4)$$

where  $\hat{\mu} = \sum_{i=1}^N \hat{B}_i / N$  is the sample mean between-subjects with  $\hat{\mu} \in \mathbb{R}^{p \times m}$ , and  $\hat{\sigma} = \frac{1}{N-1} \sum_{i=1}^N (\hat{B}_i - \hat{\mu})^2$  is the sample variance between-subjects with  $\hat{\sigma} \in \mathbb{R}^{p \times m}$ . Therefore, we have one local test  $T_k$  for each voxel/null hypothesis  $H_0^k : \mu_k = 0$  against the two-sided alternative hypothesis.

In the ProMises model (1) we assume  $M$  equals for all subjects, while in (3) the mean effect is subject-specific, that is  $M_i = DB_i = DB + DU_i = M + DU_i$ . The

additional random matrix  $DU_i$  must be then inserted in (1). However,  $\mathbb{E}(U_i) = 0$  by definition, and  $\mathbb{E}(M_i) = M$ . We assume that under  $H_0$  the random effect involved in  $B_i$  is null (i.e., these terms does not impact the estimation process of  $R_i$ ).

### 2.3 Valid Inference

Here, we propose a theorem that ensures the validity of inference using functionally aligned data.

**Theorem 1.** *Let consider the  $p$ -values  $p_k$  related to the statistical test  $T_k$ , where  $k \in M = \{1, \dots, m\}$ . If the ProMises model defined in Subsection (2.1) is valid, then:*

$$\Pr(p_k \leq \alpha \mid \hat{R}_i, \hat{\alpha}_i) = \alpha, \quad \forall k \in S \quad (5)$$

where  $S \subseteq M$  is the set of true null hypotheses,  $\alpha$  is the significance level, and  $\hat{R}_i$  defined as (2).

*Proof.* Let assume that the the ProMises model defined in Equation (1) is valid, then we have:

$$X_i = \alpha_i(M_i + E_i)R_i^\top = \alpha_i(DB_i + ZG_i + E_i)R_i^\top. \quad (6)$$

W.l.o.g. we assume the set  $\{X_i\}_{i=1, \dots, N}$  to be column-centered matrices. Write  $H_Z = Z(Z^\top Z)^{-1}Z^\top$ , so  $\hat{B}_i = (D^\top(I - H_Z)D)^{-1}D^\top(I - H_Z)X_i$ . The ProMises model becomes  $X_i = \alpha_i(ZG_i + E_i)R_i^\top$  under  $H_0$ . It is now easy to see that the  $\hat{B}_i \mid \hat{R}_i, \hat{\alpha}_i$  is normal zero-centered under  $H_0$  like  $\hat{B}_i$  with variance  $\Sigma$ . In addition, the information involved in the estimation of  $R_i$  and  $B_i$  are orthogonal (i.e.,  $\mathbb{E}((ZG_i + E_i)^\top(I - H_Z)\hat{X}_i) = \text{tr}(\Sigma_n^\top(I - H_Z))\Sigma_m^\top$  which is independent of  $D$  and other subjects's information, and  $X_{j \neq i}$  are independent of  $(I - H_Z)X_i$  by definition). This implies that also the statistical test  $T \mid \hat{R}_i, \hat{\alpha}_i$  defined in (4) is normal zero-centered and finally  $\Pr(p_k \leq \alpha \mid \hat{R}_i, \hat{\alpha}_i) = \alpha \forall k \in N$ .

Theorem 1 assures the validity of the group analysis inference when data functionally aligned by the ProMises model are used instead of the raw data. In the next section, we apply the ProMises model to fMRI images and then we perform one-sample t-tests for each voxel to analyze the group mean activation. The gain in power is notable if the functionally aligned data are used instead of the raw ones.

## 3 Application

### 3.1 Data

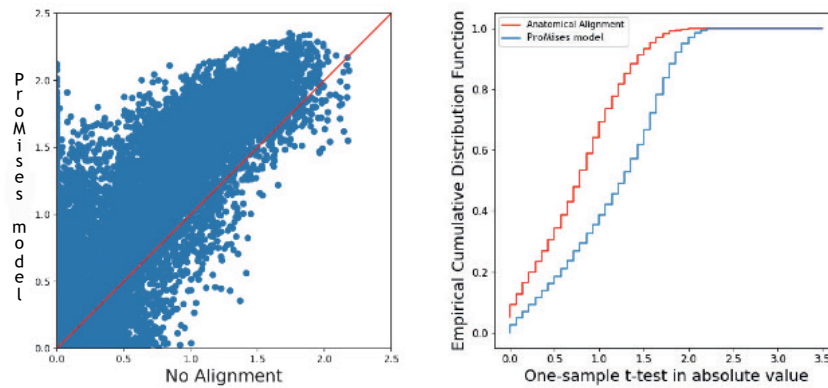
The auditory data collected by [12] are analyzed, available at <https://openneuro.org/datasets/ds000158/versions/1.0.0>. The study consists

of neural activation of 218 subjects passively listening to vocal, i.e., speech, and nonvocal sounds. We randomly select 18 subjects.

We preprocess the data using the FMRIB Software Library (FSL) [10] using a standard processing procedure. For further details about the experimental design and data acquisition, please see [12]. We perform a higher-level analysis ROI, i.e., group-subject analysis [11], considering the superior temporal gyrus (STG) as ROI, which is well known to be involved in auditory processing [6]. The STG was extracted from the Harvard Oxford cortical structural atlas (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases>). The matrices are then composed of 310 rows (time points/stimuli) and 10233 columns (voxels). The ProMises model is implemented in [3] based on the programming language Python [14] according to the PyMVPA package [9].

### 3.2 Results

Using the ProMises model, we found notable results for group analysis of fMRI data. The left panel of Figure 1 illustrates the one-sample t-tests (4) computed using the images aligned by the anatomical alignment [13] and by the ProMises model. The right panel of Figure 1 shows the empirical cumulative distribution function of these one-sample t-tests. The ProMises model returns a set of t-tests 85.85% higher in mean (in absolute value) than the t-tests computed using anatomical alignment. In the case of equivalent t-tests, the percentage would equal 50%.



**Fig. 1 Left panel:** Scatter plot of the one-sample t-tests in absolute value considering the fMRI images of 18 subjects aligned by anatomical alignment (x-axis) and the ProMises model (y-axis). **Right panel:** Empirical Cumulative Distribution Function of one-sample t-tests in absolute value considering the fMRI images of 18 subjects aligned by anatomical alignment (red line) and the ProMises model (blue line).

## 4 Discussion

In this manuscript, we evaluated the ProMises model in the context of group-level inference analysis of fMRI data. We proved the validity of the inference if functionally aligned data are used instead of the raw ones (where only the anatomical registration is performed). The information used in the ProMises model to estimate the orthogonal matrix parameter can be seen as prior information used to sort the statistical tests in the group analysis. Finally, we showed an interesting application where the one-sample t-tests at the group level using data aligned by the ProMises model illustrate higher absolute values than the one-sample t-tests computed using data aligned by the anatomical alignment.

## References

1. Basso, D. and Finos, L. (2012) Exact multivariate permutation tests for fixed effects in mixed-models. *Communications in Statistics-Theory and Methods* 41(16-17), :2991–3001.
2. Andreella, A. and Finos, L. (2022). Procrustes analysis for high-dimensional data. *Psychometrika (accepted)*. DOI: 10.1007/s11336-022-09859-5 .
3. Andreella, A. (2020). ProMises (Procrustes von Mises-Fisher) model, GitHub repository. DOI: 10.5281/zenodo.5095114.
4. Downs, D. T. (1972). Orientation statistics. *Biometrika*, 59(3):665.
5. Farcomeni, A. and Finos, L. (2013). FDR control with pseudo-gatekeeping based on a possibly data driven order of the hypotheses. *Biometrics*, 69(3):606–613.
6. Galaburda, A. and Sanides, F. (1980). Cytoarchitectonic organization of the human auditory cortex. *The Journal of Comparative Neurology*, 190(3):597–610.
7. Gower, J. C. and Dijksterhuis, B. G. (2004). *Procrustes problems*. Oxford University Press.
8. Gupta, A. K. Nagar, D. K. (2018). *Matrix variate distributions*. Chapman and Hall/CRC.
9. Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., and Pollmann, S. (2009). PyMVPA: A Python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, (7):37–53.
10. Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M. (2012). FSL. *NeuroImage*, 62:782–90.
11. Mumford, J. A. and Poldrack, R. (2007). Modeling group fMRI data. *Social Cognitive and Affective Neuroscience*, 62(2):251–257.
12. Pernet, C. R., McAleer, P. M., L., Gorgolewski, K. J., Charest, I., Bestelmeyer, P. E., Watson, R. H., Fleming, D., Crabbe, F., Valdes-Sosa, M., and Belin, P. (2015). The human voice areas: spatial organization and inter-individual variability in temporal and extra-temporal cortices. *NeuroImage*, 119:164–74.
13. Talairach, J. J. and Tournoux, P. (1988). *Co-Planar stereotaxic atlas of the human brain.3-dimensional proportional system: an approach to cerebral imaging*. Atlante.
14. Van Rossum, G. and Drake Jr, F. (1995). *Python tutorial*. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.