



HAL
open science

Quantifying the researcher degrees of freedom in FMRI preprocessing

Melvin Selim Atay, Boris Clénet, Elise Banner, Camille Maumet

► **To cite this version:**

Melvin Selim Atay, Boris Clénet, Elise Banner, Camille Maumet. Quantifying the researcher degrees of freedom in FMRI preprocessing. 2025. <hal-05365306>

HAL Id: hal-05365306

<https://inria.hal.science/hal-05365306v1>

Preprint submitted on 14 Nov 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

QUANTIFYING THE RESEARCHER DEGREES OF FREEDOM IN FMRI PREPROCESSING

Melvin Selim Atay^[1] Boris Clénet^[1] Elise Bannier^{*} [1,2] Camille Maumet^{*}[1]

^[1] Univ Rennes, Inria, CNRS, Inserm.

^[2] CHU Rennes, Radiology Department, Rennes France.

ABSTRACT

At every step of the processing of task-fMRI data, researchers have to make multiple decisions: from deciding on software and their versions to choosing parameters and their values. Such variations in the pipelines have been shown to give rise to differences in experimental results, leading to so-called *analytical variability* and hindering reproducibility.

Our aim is to investigate the task-fMRI parameter space. First we counted the number of possible preprocessing options in each software package. Second, we analyzed the parameter preferences of scientists by retrieving and interrogating open-source codes. We targeted three widely-used neuroimaging software packages: AFNI, FSL, and SPM. Using GitHub, we created a codebase containing a total of 3000 code files from 93, 157, and 312 unique repositories for AFNI, FSL, and SPM, respectively.

We observe fundamental differences between software in terms of pipeline steps, number of parameters available, and parameter values chosen by researchers. We detected variations in the default values per parameter for each software as well as different researcher preferences.

In the codebase, the most used values for smoothing kernel parameter vary. For AFNI 45% of repositories choose the non default kernel value (6 mm), while for FSL 70% of repositories use the default kernel value (5 mm) and, 40% of SPM repositories use the default values (8 mm).

Index Terms— Analytical variability, reproducibility, source code analysis, open source, task-fMRI.

1. INTRODUCTION

Neuroimaging pipelines, specifically analysis steps of task functional Magnetic Resonance Imaging (task-fMRI) data, require the use of multi-step complex algorithms orchestrated by different software packages. Every time a researcher constructs a pipeline, they face multidimensional choices in deciding on software, software versions, pipeline steps. The process of making multiple decisions amongst these choices is termed the *researcher degrees of freedom*.

For each step, there are various parameter choices and multiple different values, which we define as *parameter space*. The process of deciding on parameter values is denoted as *parameter preferences*. As an effect of parameter preferences of researchers, neuroimaging pipelines and subsequent results may vary. The differences observed in the pipelines are later referred to as *pipeline variations*.

In order to better understand which pipelines are used *in the wild*, various authors have looked at methods reporting in the task-fMRI literature. In their survey [1], Carp demonstrated variations in the reporting of research results. They highlighted reporting differences in papers at several stages of the fMRI data life cycle, including experimental design, acquisition, preprocessing, and statistical analysis.

But previous literature has also shown how typical pipeline variations can have a substantial impact on final results and hence hinder reproducibility. In their landmark paper [2], Botvinik-Nezer et al. proposed a collaborative study to investigate how different pipelines can lead to different results in neuroimaging. They shared a dataset along with nine pre-defined hypotheses to be tested. Seventy teams took part in the challenge and independently analyzed the dataset to answer the research hypotheses. Each team described the pipeline they used. Overall differences in pipelines sometimes led to contradictory conclusions.

Furthermore, to assess the impact of software packages on the robustness of the results, Bowring et al. inspected the preprocessing and analysis steps for AFNI, FSL, and SPM pipelines on three different task-fMRI studies [3]. They emphasized the variability induced by different neuroimaging software packages in the results of higher-level analyzes.

To better harmonize and increase the level of detail of the Methods sections, there have been efforts to build guidelines for method reporting [4]. Other efforts focused on facilitating and encouraging code sharing alongside research papers [5]. Increasing code sharing effectively creates new opportunities to better understand pipelines by directly examining the source code accompanying scientific papers.

Yet, it remains unknown what are the main levels of pipeline variations that impact analytical variability. A better understanding of the researcher degrees of freedom in neuroimaging pipelines is therefore needed to investigate the

* Contributed equally.

sources of analytical variability.

Here, we investigate the researcher degree of freedom in the task-fMRI preprocessing pipelines by delving into open-source scripts of neuroimaging analyzing pipelines to reveal pipeline variations. The collection of corresponding code files are referred to as pipeline scripts for simplification. In the following sections, these open source neuroimaging analysis pipeline scripts are shortened as *pipeline scripts*.

Henceforth, we focus on three of the most common task-fMRI software packages: Analysis of Functional NeuroImages (AFNI) [6], the FMRIB Software Library (FSL) [7], and Statistical Parametric Mapping (SPM) [8].

2. METHODS

2.1. Overview of the parameter space

To quantify how many options are available to researchers, we listed all the parameters available on the user interfaces of `afni_proc` [9] (AFNI), FEAT (FSL) and `slice timing`, `Realign`, `Coregister`, `Segment Normalize`, and `Smooth` (SPM).

We further classified the parameters into three groups: 1) acquisition parameters, i.e., parameters that depend on data acquisition or experimental setup (e.g., value of the TR, acquisition resolution); 2), data input/output parameters, i.e., parameters connecting different sub-steps of the pipeline (e.g. anatomical MRI as an input of segmentation) and 3) analysis choices (e.g., interpolation method).

Repeated instances of parameters were counted once. As an example, the 6 parameters for each of the 6 tissue probability maps defined in SPM for the segmentation step were counted as 1 parameter.

We chose to focus on parameters of the preprocessing steps, and not parameters of the statistical analysis that are more tight to the research question and the experimental design. To keep only preprocessing pipelines, we filtered out the pipeline scripts that did not contain any motion correction step.

2.2. Analysis of pipeline variations

Our code search and analysis methodology is a two step approach (see Fig. 1). First, we designed and run the query to retrieve pipeline scripts. Second, we ran analyzes on the codebase to estimate the parameter preferences across the parameter space.

2.2.1. Creation of the Codebase of Pipeline Scripts

To form our codebase, we designed unique queries to match the AFNI, FSL, and SPM pipeline scripts from GitHub, based on the rules described in Table 1. We implemented queries with GitHub REST API (v3) and PyGitHub (version 2.7.0). Since the GitHub search functionality is limited to returning

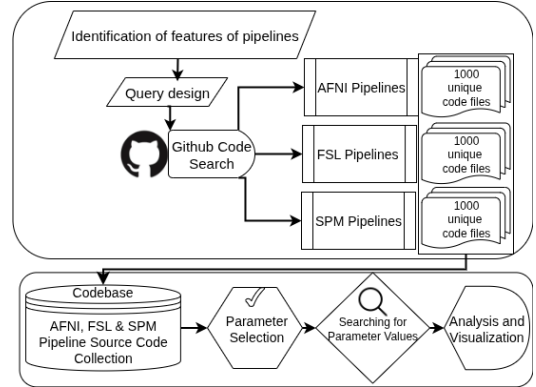


Fig. 1: Summary of data collection and analysis steps. 1. Query : Defining unique search parameters for pipelines. 2. Codebase and Analysis: Parameter value search and visualizations on the collection of repositories.

Software	File Contains	Extension
SPM	<code>spm.stats_fmri_spec</code>	<code>.m</code>
FSL	<code>set_fmri(version)</code>	<code>.fsf</code>
AFNI	<code>afni_proc.py</code>	(no filtering)

Table 1: Search phrases used for querying pipeline scripts.

1000 code files per query, we performed a separate query per software to maximize the number of results returned. The queries were run on June 03, 2025.

As can be inferred from the query rules, we included only task-fMRI pipelines, excluded AFNI, FSL, and SPM source codes, as well as animal pipelines (exclusion of repositories containing animal names such as *rat*, *mice*, *monkey*, *pig*, *sheep* and *marmoset* in the folder names or names of the repositories.).

2.2.2. Assessment of Parameter Preferences

We mapped the correspondence for each parameter of the user interfaces (see 3.1) to its representations in the code. We extracted the parameters in the pipelines along with their values. As an example, the full width at half maximum of the spatial smoothing kernel in `afni_proc.py` is given after the flag `-blur_size`. In `design.fsf` in FSL, it is defined after `set_fmri(smooth)`. And in SPM batch files as the value of `spm.spatial.smooth.fwhm`.

3. RESULTS AND DISCUSSION

3.1. Overview of the parameter space

We present the number of parameters available for each pre-defined category (researchers' choice, Input/Output, and acquisition) in Table 2. The number of parameters chosen by researchers varies between software packages with a total of 45 parameters for SPM, 29 for FSL and, 40 for AFNI. As a note, the segmentation parameters in AFNI are listed as NA since `afni_proc.py` does not include a segmentation step [6].

Processes	Software								
	SPM			FSL			AFNI		
	Choice	I/O	Acq.	Choice	I/O	Acq.	Choice	I/O	Acq.
Slice Timing	1	2	4	1	1	1	3	1	1
Motion Correction	11	3	0	1	0	0	8	1	3
Coregistration	10	0	0	6	0	0	9	2	1
Segmentation	12	3	0	13	2	0	NA	NA	NA
Spatial Normalization	9	3	0	7	2	0	17	3	0
Spatial Smoothing	2	3	0	1	0	0	3	0	0
TOTAL	45	14	4	29	5	1	40	7	5

Table 2: Overview of the parameter space: number of parameters for each preprocessing step. (Choice= Chosen by researchers; I/O= Data Input-Outputs, Acq.= Acquisition-dependent).

3.2. Analysis of pipeline variability

3.2.1. Codebase of pipelines

The codebase is composed of 3000 code files coming from 93, 157, and 312 repositories in AFNI, FSL, and SPM, respectively. These repositories correspond to various preprocessing steps and cover a period from 2010 to 2025.

Fig. 2 shows the percentage of repositories for each software package since 2016. (We excluded earlier dates due to a low study count per year per software, $N < 8$). SPM corresponds to the highest ratio followed by FSL and AFNI. We note that this distribution is consistent with prior work based on reports found in the literature [1].

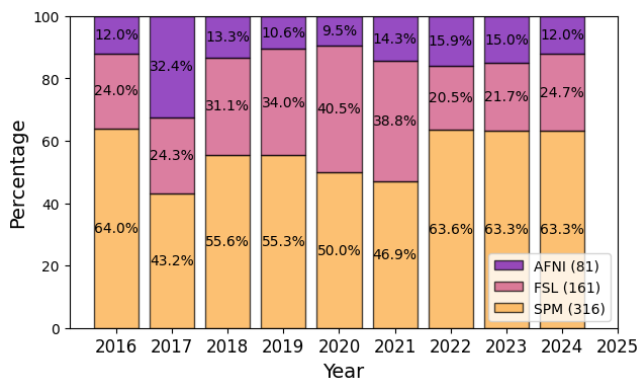


Fig. 2: Percentage of repositories for each software per year. SPM pipeline scripts have the highest number of repositories.

3.2.2. Analysis of the use of preprocessing steps

The use (or not) of each preprocessing step is summarized in Fig. 3. In FSL, nearly all preprocessing steps are always implemented, except for the slice timing correction (present in 19% of the repositories) and the segmentation (96%). For SPM pipeline scripts, the least implemented step is slice timing correction (67%), followed by normalization (80%), coregistration, and segmentation are the most used steps (both

93%). In AFNI pipeline scripts there are more variations, the most performed step is coregistration (87%) and the least is slice timing (34%).

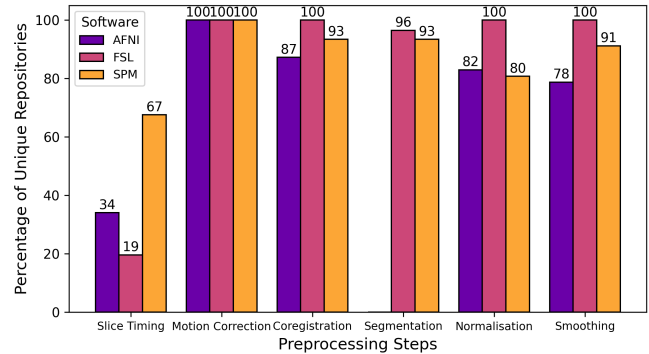


Fig. 3: Comparison of the use of preprocessing steps. Across all software, coregistration is the most implemented whereas slice timing correction is the least used.

3.2.3. Assessment of Parameter Preferences

As emphasized above, pipeline steps vary between software. To further illustrate some of the differences in parameter value choices, we looked more closely at the parameters for the smoothing and motion correction steps. The values taken by the smoothing kernel are shown in Fig. 4.

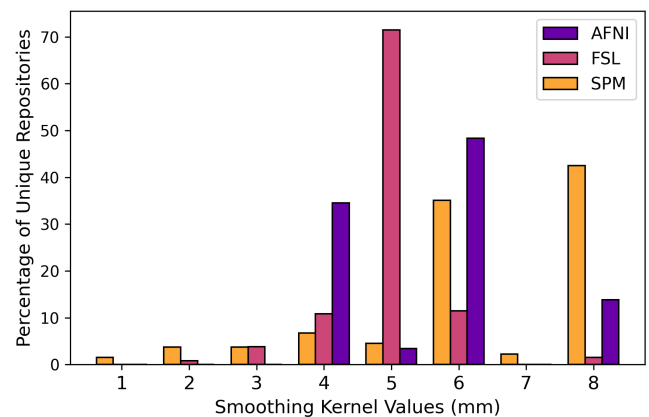


Fig. 4: Parameter value choices for smoothing kernel. Distribution of smoothing kernel values for each software pipeline is varying.

In AFNI, the most commonly used smoothing kernel value is 6 mm (45%) followed by the default smoothing kernel value of 4 mm (35%).

In both FSL and SPM, the default smoothing kernel value is the most used: 5 mm in FSL (70%) and 8 mm in SPM (40%). In SPM, the second most represented value is 6 mm (35%). Those variations may be related to acquisition resolution, although verification of these values is beyond the reach of pipeline scripts. In Fig. 5, we illustrate the varying value

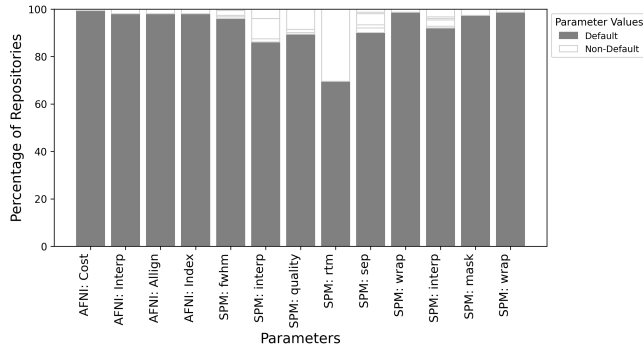


Fig. 5: Motion correction parameter values in SPM and AFNI. AFNI:Cost=cost function, AFNI:interp=interpolation method, AFNI:Allign=Extra cross-run registration, AFNI:Index=Per run index for Allign option. SPM:interp=degree of the interpolation spline for estimation, SPM:quality=quality of realignment SPM:rtm=Register to mean, SPM:sep=Separation between points sampled SPM:interp=degree of interpolation spline for writing SPM:mask=Masking for motion SPM:wrap direction of images to wrap.

choices of SPM and AFNI motion correction parameters.

For SPM pipelines, parameter with the most variation is "registering to the mean or the first slice" (parameter `eoptions.rtm`). There, 40% of the pipelines choose to register to the first slice instead of the mean (the default). The default value is based on the PET data settings. In MRI it is recommended to select the value of `eoptions.rtm` to enable registration to the first image [8]. Although this difference is documented in the manual [10] page 28, researchers tend to keep the default value defined by the software.

AFNI pipelines have more uniformity, 99% of cost function parameter take the default value, and similarly for the rest of the parameters take the software default values (98%).

4. CONCLUSION

We explored the parameter space and the pipeline variations chosen by researchers in 3 neuroimaging software packages. The different practices observed for slice timing correction are in parallel with community discussions that point to the optionality of this step [11].

In our codebase, we observed varying parameter choices. For motion correction parameters, the default values of the software are most often selected. The reason behind this choice may be to ease the complexity of algorithmic decisions.

This work is based on a limited number of pipelines collected from GitHub code search. In the future, to perform a more exhaustive analysis, we will work on a larger collection of source code scripts.

5. ACKNOWLEDGMENTS

This work was supported by the National Research Agency under the France 2030 program [ANR-22-PESN-0007].

References

- [1] J. Carp, "The secret lives of experiments: Methods reporting in the fMRI literature," *NeuroImage*, vol. 63, no. 1, pp. 289–300, Oct. 2012, publisher: Elsevier BV.
- [2] R. Botvinik-Nezer, F. Holzmeister, C. F. Camerer, A. Dreber, J. Huber, M. Johannesson *et al.*, "Variability in the analysis of a single neuroimaging dataset by many teams," *Nature*, vol. 582, no. 7810, pp. 84–88, Jun. 2020.
- [3] A. Bowring, T. E. Nichols, and C. Maumet, "Isolating the sources of pipeline-variability in group-level task fmri results," *Human Brain Mapping*, vol. 43, no. 3, pp. 1112–1128, Feb. 2022.
- [4] T. E. Nichols, S. Das, S. B. Eickhoff, A. C. Evans, T. Glatard, M. Hanke *et al.*, "Best practices in data analysis and sharing in neuroimaging using mri," *Nat. Neurosci*, vol. 20, no. 3, pp. 299–303, 2017.
- [5] R. A. Poldrack and K. J. Gorgolewski, "Making big data open: data sharing in neuroimaging," *Nat. Neurosci*, vol. 17, no. 11, pp. 1510–1517, Nov. 2014.
- [6] R. W. Cox, "AFNI: What a long strange trip it's been," *NeuroImage*, vol. 62, no. 2, pp. 743–747, Aug. 2012.
- [7] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *NeuroImage*, vol. 62, no. 2, pp. 782–790, Aug. 2012.
- [8] K. J. Friston, "Statistical Parametric Mapping," in *Neuroscience Databases*. Boston, MA: Springer US, 2003, pp. 237–250.
- [9] R. C. Reynolds, D. R. Glen, G. Chen, Z. S. Saad, R. W. Cox, and P. A. Taylor, "Processing, evaluating, and understanding FMRI data with `afni.proc.py`," *Imaging Neuroscience*, vol. 2, pp. imag-2-00 347, Nov. 2024.
- [10] J. Ashburner, G. Barnes, C.-C. Chen, J. Daunizeau, G. Flandin, K. Friston *et al.*, "Spm12 manual," *Wellcome Trust Centre for Neuroimaging, London, UK*, vol. 2464, no. 4, p. 53, 2014.
- [11] R. Sladky, K. J. Friston, J. Tröstl, R. Cunnington, E. Moser, and C. Windischberger, "Slice-timing effects and their correction in functional MRI," *NeuroImage*, vol. 58, no. 2, pp. 588–594, Sep. 2011.