

# How does group differences in motion scrubbing affect false positives in functional connectivity studies?

Anders Eklund<sup>a,b,c</sup>, Thomas E. Nichols<sup>d,e,f</sup>, Soroosh Afyouni<sup>d</sup>, Cameron Craddock<sup>g</sup>

<sup>a</sup>*Division of Medical Informatics, Department of Biomedical Engineering, Linköping University, Linköping, Sweden*

<sup>b</sup>*Division of Statistics & Machine Learning, Department of Computer and Information Science, Linköping University, Linköping, Sweden*

<sup>c</sup>*Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden*

<sup>d</sup>*Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, UK*

<sup>e</sup>*Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, OX3 9DU, UK*

<sup>f</sup>*Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK*

<sup>g</sup>*Department of Diagnostic Medicine, Dell Medical School, Austin, Texas, USA*

---

## Abstract

Analyzing resting state fMRI data is difficult due to a weak signal and several noise sources. Head motion is also a major problem and it is common to apply motion scrubbing, i.e. to remove time points where a subject has moved more than some pre-defined motion threshold. A problem arises if one cohort on average moves more than another, since the remaining temporal degrees of freedom are then different for the two groups. The effect of this is that the uncertainty of the functional connectivity estimates (e.g. Pearson correlations) are different for the two groups, but this is seldom modelled in resting state fMRI. We demonstrate that group differences in motion scrubbing can result in inflated false positives, depending on how the temporal auto correlation is modelled when performing the Fisher r-to-z transform.

*Keywords:* fMRI, motion scrubbing, false positives

---

## 1. Introduction

Motion scrubbing is commonly applied in resting state fMRI, to remove volumes where the head motion is greater than some threshold (Power et al., 2012). Motion scrubbing is performed independently for every subject, meaning that some subjects can have 90% of the fMRI volumes left after motion scrubbing, while other subjects may only have 50% of the volumes left. As an example, in the paper by Power et al. (2012) it is reported that  $58\% \pm 20\%$  and  $26\% \pm 14\%$  of the volumes were removed for two groups of children, but only  $12\% \pm 6\%$  of the volumes were removed for a group of adults. Siegel et al. (2014) reported  $4\% \pm 5\%$  scrubbing for typical adults and  $16\% \pm 11\%$  scrubbing for children. Similarly, Parkes et al. (2018) reported a significant difference in loss of temporal degrees of freedom between healthy controls and schizophrenics.

Group differences in motion scrubbing means that the average uncertainty of the functional connectivity estimates (e.g. Pearson correlations) can differ between the two groups, since the estimates are based on different number of measurements. Propagating subject uncertainty to the group analysis is, however, uncommon in resting state fMRI (an exception is the work by Fiecas et al. (2017)). This is in

contrast to task fMRI, where the variance of each subject is commonly used in the group analysis (Woolrich et al., 2004; Chen et al., 2012), to for example downweight subjects with a higher variance. The problem of variable temporal degrees of freedom in resting state fMRI has been mentioned previously (Yan et al., 2013; Satterthwaite et al., 2013; Power et al., 2014; Pruim et al., 2015; Parkes et al., 2018), but to the best of our knowledge no study has investigated how it affects false positives in a group analysis. Here, we demonstrate that group differences in motion scrubbing can lead to inflated false positives rates for functional connectivity studies.

## 2. Data

We used the ABIDE preprocessed dataset (Di Martino et al., 2014; Craddock et al., 2013a)<sup>1</sup>, which contains preprocessed resting state fMRI data from 539 individuals diagnosed with autism spectrum disorder and 573 typical controls. The data were preprocessed using four different preprocessing pipelines (Connectome Computation System (CCS) (Xu et al., 2015), Configurable Pipeline for the Analysis of Connectomes (CPAC) (Craddock et al., 2013b), Data Processing Assistant for Resting-State fMRI (DPARSF) (Yan & Zang, 2010), NeuroImaging Analysis

---

\*Corresponding author

Email address: [anders.eklund@liu.se](mailto:anders.eklund@liu.se) (Anders Eklund)

<sup>1</sup><http://preprocessed-connectomes-project.org/abide/>

Kit (NIAK) (Bellec et al., 2011)), and are available for several preprocessing options (such as with and without bandpass filtering, and with and without global signal regression (GSR)<sup>2</sup>). Motion scrubbing has been applied for the NIAK pipeline, but not for the other pipelines, and we therefore used preprocessed data from CCS, CPAC and DPARSF. Mean timeseries were extracted for 7 regions of interest (ROI) atlases, and we focused our analyses on the Craddock 200 (CC200) atlas (Craddock et al., 2012), which consists of 200 ROIs. To lower the processing time, and to improve partial correlation estimates, we only performed the calculations for 50 of the 200 ROIs.

The ABIDE data were collected at 17 different sites, and all group analyses reported in this paper only used data from one site at a time. We focused on data collected at Michigan and New York, as these datasets contain the highest number of subjects (110 and 184, respectively). The downloaded data contains 82 subjects for Michigan and 171 subjects for New York, which is mainly explained by the fact that subjects with a mean frame-wise displacement larger than 0.2 are by default discarded by the ABIDE preprocessed download script. The Michigan datasets contain 296 volumes per subject, while the New York datasets contain 176 volumes per subject. All datasets were collected with a TR of 2 seconds.

### 3. Methods

#### 3.1. Motion scrubbing

To investigate the effect of group differences in motion scrubbing on functional connectivity analyses, we used the same idea as in our previous work (Eklund et al., 2016, 2019); to perform many random group analyses with real fMRI data to empirically estimate false positives. Two random groups, of 20 subjects each, were first created by randomly shuffling all subjects and selecting the 40 first subjects. Preprocessed ABIDE data were then randomly motion scrubbed for every subject, where the proportion of scrubbing for each subject was randomly drawn from a normal distribution (at least 50 volumes were always saved). The mean proportion of scrubbing for the (fake) control group was set to 10%. For the (fake) diseased group, the mean proportion of scrubbing varied from 5% to 50%, in steps of 5% units. The standard deviation was set to 5% for the control group and 15% for the diseased group (Power et al., 2012). To reflect the fact that motion contaminated volumes often appear in clusters, rather than appearing randomly, a simulation script was used to create more realistic motion scrubbing, see Figure 1.

#### 3.2. ROI to ROI correlation estimation

Pairwise correlations (both full and partial) between all ROI time series were calculated using the motion scrubbed data (using the function `nets_netmats` in FSLNets<sup>3</sup> (Smith

et al., 2011)). For partial correlation we used L2 norm regularization (Tikhonov), since L1 regularization is rather slow, and the default setting in FSLNets (0.1). In resting state fMRI the correlation values are normally Fisher r-to-z transformed before applying a t-test. Note that the Fisher transformation on its own will not produce z-scores, since it does not involve division with a standard error. The temporal auto correlation present in fMRI data complicates the variance estimation (Eklund et al., 2012; Arbabshirani et al., 2014; Afyouni et al., 2019), which can lead to inflated z-scores. FSLNets uses a global AR(1) model (the same for all ROI to ROI correlations) to model the temporal auto correlation. The AR parameter is estimated using a Monte Carlo simulation (and not through prewhitening (Bright et al., 2017; Honari et al., 2019)). Other softwares (such as REST (Song et al., 2011) and Conn (Whitfield-Gabrieli & Nieto-Castanon, 2012)) do not mention any model for the auto correlation. We therefore decided to perform the analyses with uncorrected (raw) and corrected (FSLNets global AR(1)) r-to-z transformation.

#### 3.3. Group analysis

The group analysis consisted of applying a two sample t-test (using `ttest2` in Matlab) to every element below the diagonal in the (r-to-z transformed) correlation matrix, to test if there is a difference in mean functional connectivity between the two randomly created groups (controls > diseased). Bonferroni correction was used to correct for multiple comparisons over correlation matrix elements, and should be conservative since the tests are not independent. A total of 96 different parameter combinations, listed in Table 1, were tested to understand how different parameters affect the results. The whole procedure was repeated 1000 times for every parameter combination, to empirically estimate the (familywise error corrected) degree of false positives (defined as the proportion of times a significant group difference was found, after correcting for multiple comparisons).

#### 3.4. Reproducibility

As for our previous work (Eklund et al., 2015, 2016, 2019), our results are fully reproducible since we used open data and share the processing scripts on GitHub<sup>4</sup>. Other researchers can thereby reproduce our findings and extend the analyses to other settings (Eklund et al., 2017b; Poldrack et al., 2017).

<sup>2</sup><http://preprocessed-connectomes-project.org/abide/Pipelines.html>

<sup>3</sup><https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLNets>

<sup>4</sup><https://github.com/wanderine/GroupDifferencesMotionScrubbing>

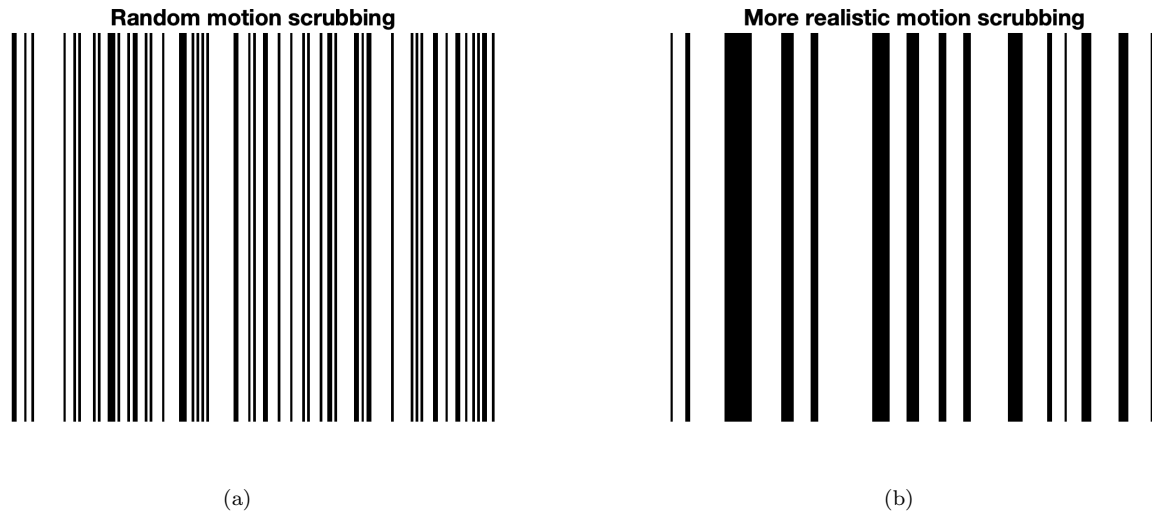


Figure 1: Comparing random scrubbing (left, obtained by the function `randperm` in Matlab) and more realistic scrubbing (right, obtained through a more advanced simulation) for a time series with 200 time points and 30% scrubbing (black time points represent time points to be removed).

Table 1: Parameters tested for the different processing pipelines. One thousand random group analyses were performed for each parameter combination.

Parameter	Values used
fMRI data	New York (171 subjects), Michigan (82 subjects)
Pipeline	CCS, CPAC, DPARSF
Bandpass filtering (0.01 - 0.1 Hz)	No, Yes
Global signal regression	No, Yes
Fisher r-to-z transform	Without autocorrelation correction (raw), With autocorrelation correction (AR(1))
Analysis type	Full correlation, partial correlation (Tikhonov regularization)

#### 4. Results

Figure 2 shows estimated FWE rates for different amounts of motion scrubbing per group and different preprocessing strategies, for the CCS pipeline and the New York data. Figures 3 and 4 show corresponding FWE rates for the CPAC and DPARSF processing pipelines. Figures 5 to 7 show corresponding results for the Michigan data. Clearly, using the r-to-z transform with auto correlation correction in FSLNets leads to inflated false positive rates for large group differences in motion scrubbing. Applying no correction for auto correlation leads to nominal results for all parameter combinations, and this is explained by the fact that the correlation (and the raw Fisher transformed correlation) are not test statistics, since they have not been divided by a standard error.

The FWE rate is in general lower when bandpass filtering is applied (mean 7.3%, compared to 12.5% without filtering). Global signal regression has a small effect on the mean FWE rate; 9.7% with GSR and 10.1% without GSR. The FWE rate is slightly higher for the New York data (mean 11.1%, compared to 8.8% for Michigan), which is expected since the Michigan data contains 296 time points per subject, compared to 176 time points

per subject for the New York data. Furthermore, the FWE rate is in general higher for partial correlation (mean 12.3%, compared to 7.5% for full correlation). The mean FWE rate per pipeline is 5.7% (CCS), 12.5% (CPAC) and 11.5% (DPARSF). A possible explanation is that the CCS pipeline includes a despiking step, which will reduce ROI correlations due to spikes (as a spike will be present in most ROIs).

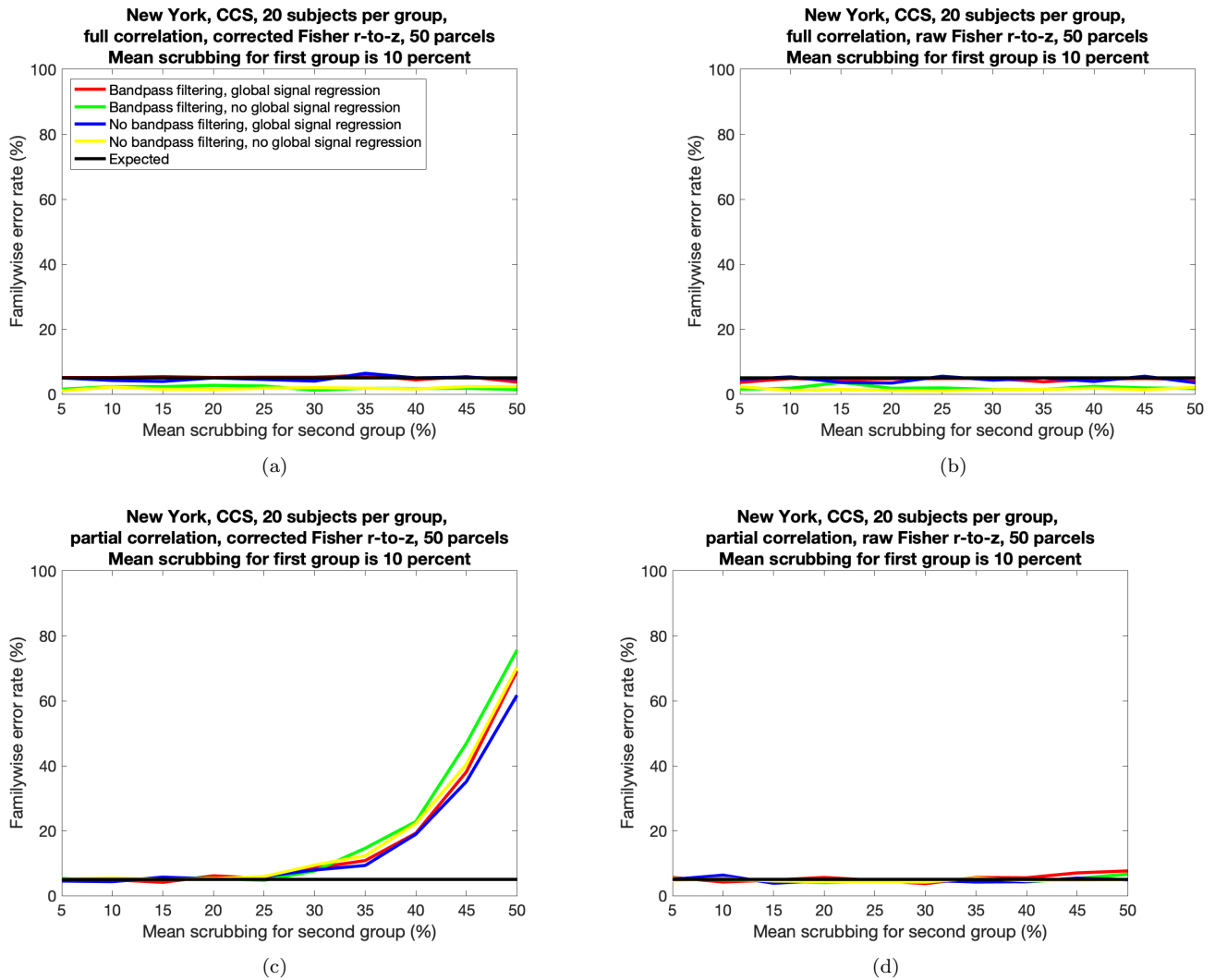


Figure 2: *Estimated familywise error rates, for different proportions of motion scrubbing for two groups of randomly selected subjects, for the CCS processing pipeline. Left: Results with Fisher r-to-z transform with auto correlation correction Right: Results with Fisher r-to-z transform without auto correlation correction Top: Results for full correlation. Bottom: Results for partial correlation.*

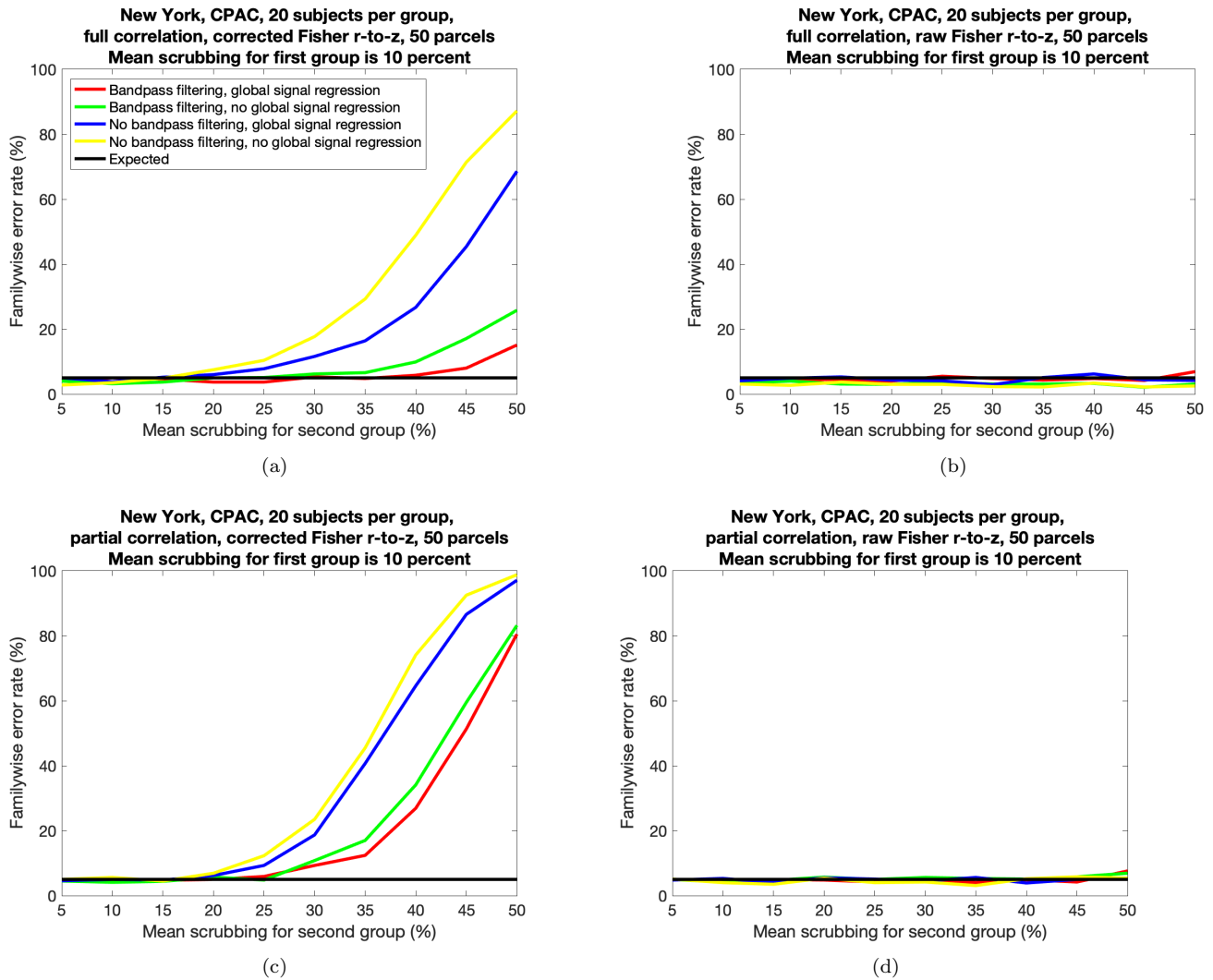


Figure 3: *Estimated familywise error rates, for different proportions of motion scrubbing for two groups of randomly selected subjects, for the CPAC processing pipeline. Left: Results with Fisher r-to-z transform with auto correlation correction Right: Results with Fisher r-to-z transform without auto correlation correction Top: Results for full correlation. Bottom: Results for partial correlation.*

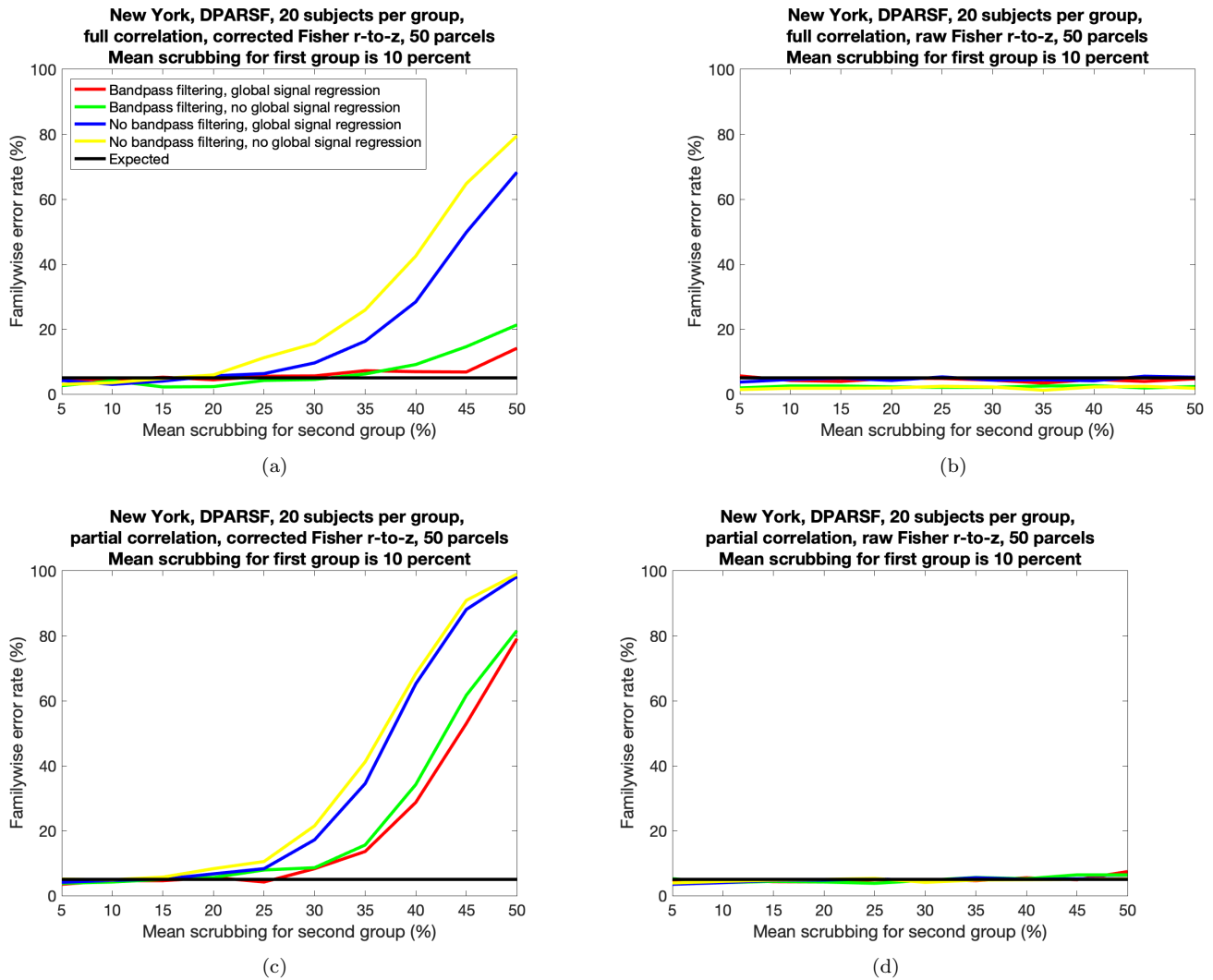


Figure 4: *Estimated familywise error rates, for different proportions of motion scrubbing for two groups of randomly selected subjects, for the DPARSF processing pipeline. Left: Results with Fisher r-to-z transform with auto correlation correction. Right: Results with Fisher r-to-z transform without auto correlation correction. Top: Results for full correlation. Bottom: Results for partial correlation.*

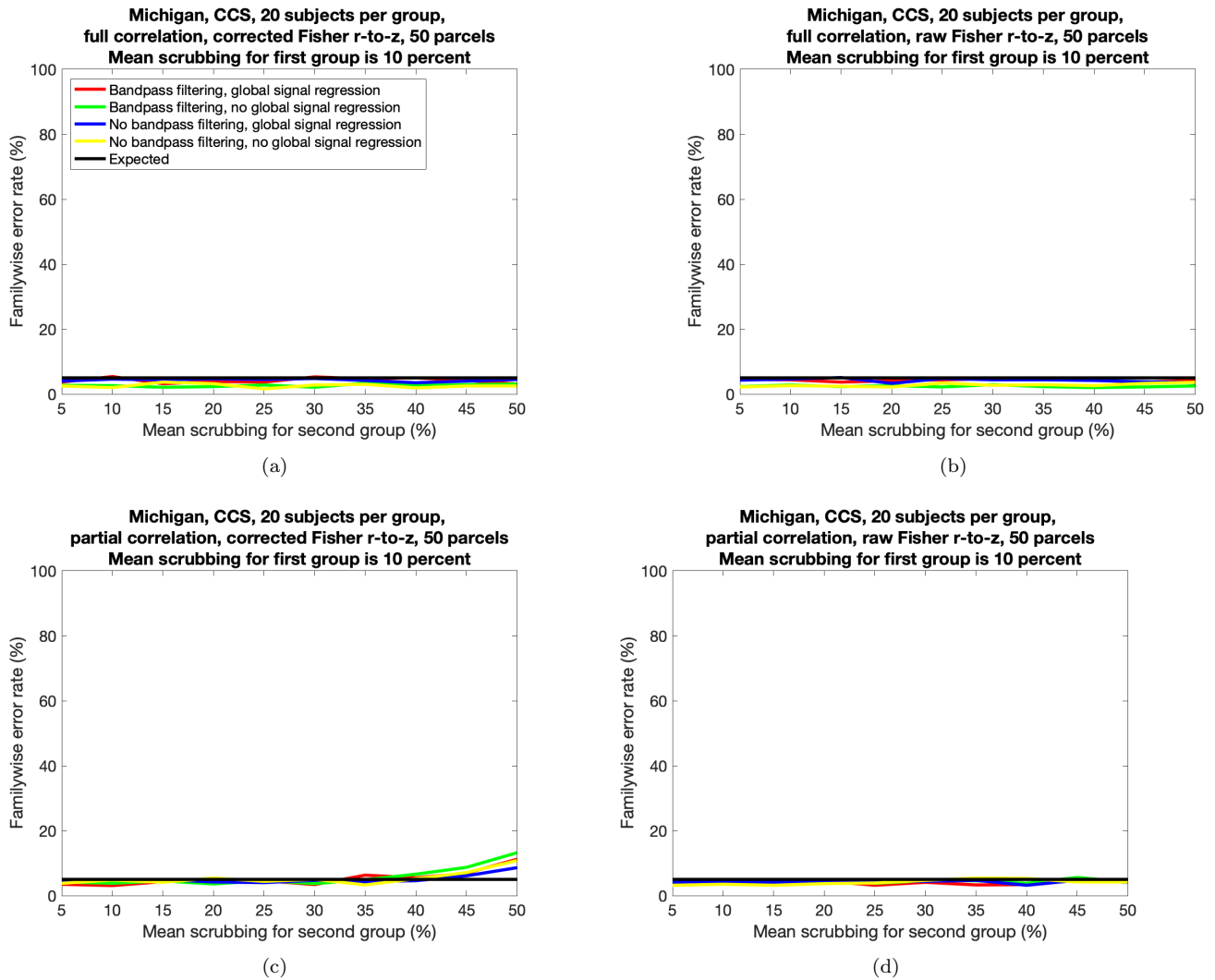


Figure 5: *Estimated familywise error rates, for different proportions of motion scrubbing for two groups of randomly selected subjects, for the CCS processing pipeline. Left: Results with Fisher r-to-z transform with auto correlation correction Right: Results with Fisher r-to-z transform without auto correlation correction Top: Results for full correlation. Bottom: Results for partial correlation.*



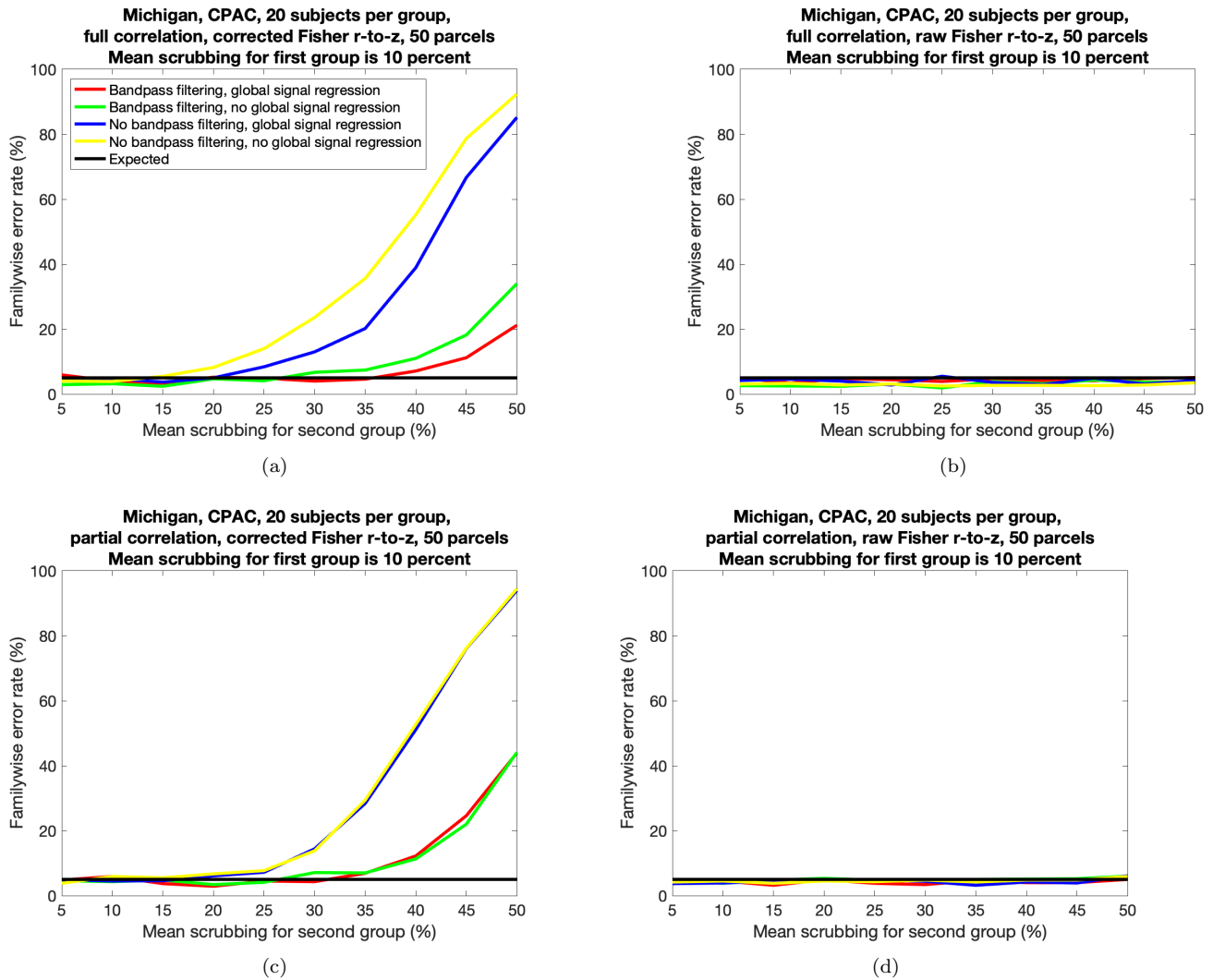


Figure 6: *Estimated familywise error rates, for different proportions of motion scrubbing for two groups of randomly selected subjects, for the CPAC processing pipeline. Left: Results with Fisher r-to-z transform with auto correlation correction Right: Results with Fisher r-to-z transform without auto correlation correction Top: Results for full correlation. Bottom: Results for partial correlation.*

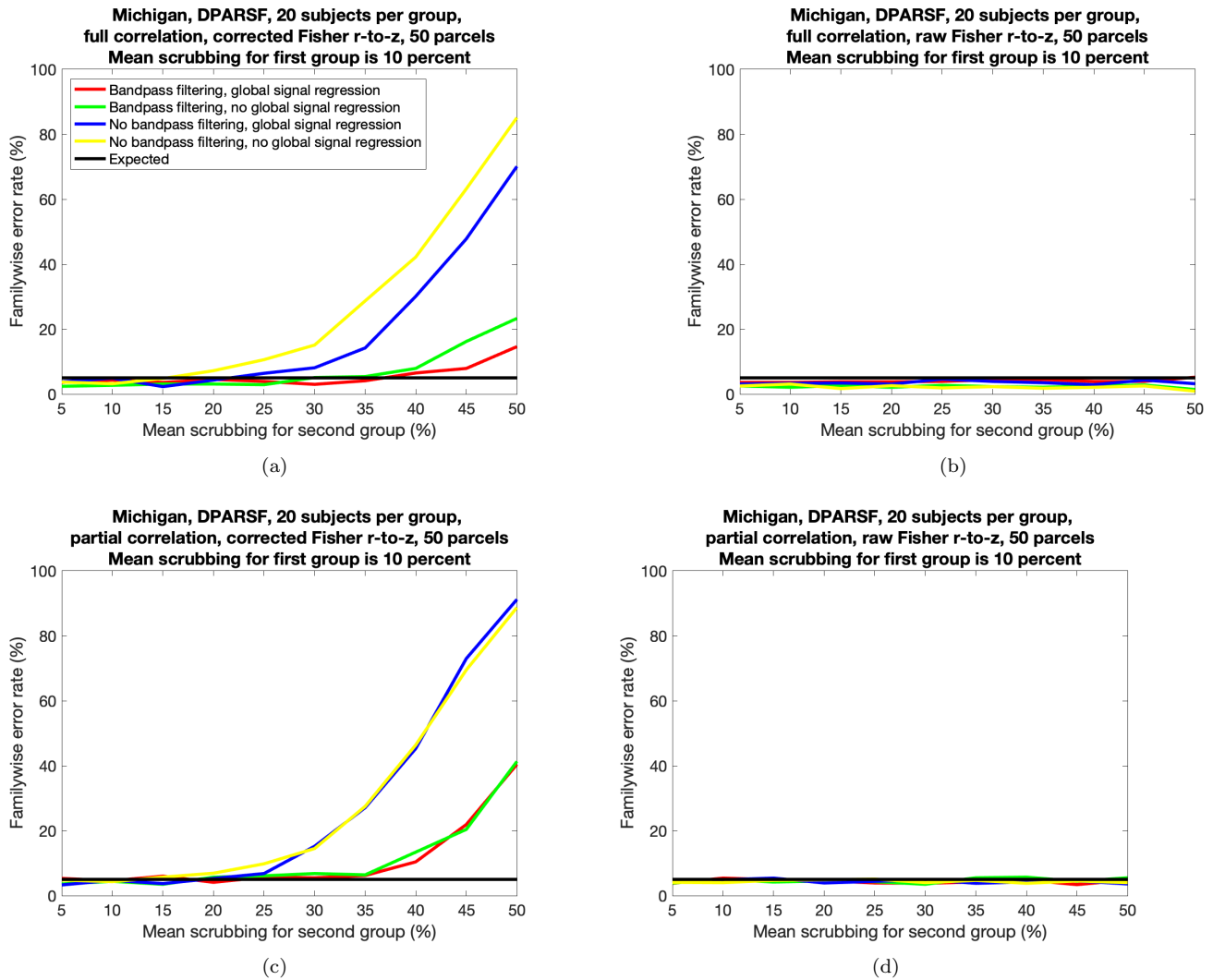


Figure 7: Estimated familywise error rates, for different proportions of motion scrubbing for two groups of randomly selected subjects, for the DPARSF processing pipeline. **Left:** Results with Fisher r-to-z transform with auto correlation correction **Right:** Results with Fisher r-to-z transform without auto correlation correction **Top:** Results for full correlation. **Bottom:** Results for partial correlation.

## 5. Discussion

We have presented results showing that group differences in motion scrubbing can lead to inflated false positive rates for functional connectivity studies, depending on how the r-to-z transform is performed. The problem of varying temporal degrees of freedom has been mentioned previously (Yan et al., 2013; Satterthwaite et al., 2013; Power et al., 2014; Pruim et al., 2015; Parkes et al., 2018), but we are not aware of any study that investigates how it affects false positives. Yendiki et al. (2014) investigated how group differences in motion affect diffusion MRI results, but did not use motion scrubbing or censoring. We have here focused on scrubbing (Power et al., 2012), but the results may also apply for spike regression (Satterthwaite et al., 2013) since adding one covariate per censored volume will also reduce the temporal degrees of freedom. It is not clear if scrubbing and spike regression would have the same effect on the autocorrelation structure of the time series. Instead of completely removing time points with too large motion, a heteroscedastic model can be used to automatically downweight affected time points (Eklund et al., 2017a). As in task fMRI, the uncertainty of each subject should also be propagated to the group analysis (Woolrich et al., 2004; Chen et al., 2012; Fiecas et al., 2017), to downweight subjects with a higher uncertainty.

### 5.1. False positives

Our results show that group differences in motion scrubbing only leads to inflated false positive rates if the z-score calculation in FSLNets accounts for the temporal auto correlation. As mentioned in the Methods section, the reason for this is that the correlation (and raw Fisher transformed correlation) are not test statistics, since they have not been divided by a standard error. The raw Fisher r-to-z transform, used in REST (Song et al., 2011) and Conn (Whitfield-Gabrieli & Nieto-Castanon, 2012), does not consider the number of time points in the time series, and will therefore produce the same z-score for a certain Pearson correlation, regardless if the number of time points used to estimate the correlation is 50 or 500 (and regardless of the temporal auto correlation in the data). As for task fMRI, where unbiased beta coefficients from the first level analyses are used in the group analysis (and not test statistics like t- or z-scores), the correlation as well as the raw Fisher transformed correlation will be (approximately) unbiased and will therefore lead to nominal results (even for large group differences in motion scrubbing).

The mean FWE rate for CCS is lower compared to CPAC and DPARSF. A possible explanation is that the CCS pipeline includes a despiking step, which will reduce ROI correlations due to spikes (as a spike will be present in most ROIs).

### 5.2. Limitations

The motion scrubbing used in this paper is completely random, which means that low motion time points can be

removed while high motion time points are still present in the data. The reason for using this approach is that the ABIDE preprocessed dataset does not contain motion metrics for each time point. In future work we will perform more realistic motion scrubbing, based on different motion metrics, such that high motion time points are the most likely to be removed.

The group analyses are based on using data from both individuals diagnosed with autism spectrum disorder and typical controls. Ideally the group analyses should be performed using only controls, but this will lead to a small number of subjects from each site.

We have here only looked at false positives, but false negatives are also important (Noble et al., 2020). To estimate statistical power using real fMRI data is not a trivial task. For group analyses involving a single group, the brain activity or connectivity of a large cohort of subjects (e.g. 400) can be seen as the ground truth, and power can then be estimated as the number of times a group analysis with a smaller number of subjects (e.g. 20) results in the same significant activity or connectivity (Lohmann et al., 2018). For group analyses involving two groups of subjects (our case), it is necessary to start with an analysis that provides a significant difference between the two groups, or to artificially add a difference and count how many times it is detected (Dansereau et al., 2017). In future work we will investigate how this can be done for resting state fMRI.

In the function `nets_glm` in FSLNets a permutation test is used to perform the group analysis and to obtain corrected p-values. We intend to try the permutation based approach, to see if it reduces the degree of false positives, but without GPU acceleration it is rather time consuming to use for so many group analyses and parameter combinations. We will investigate if BROCCOLI (Eklund et al., 2014) can be used instead of `randomise`, to speedup the permutations. Furthermore, a bug was detected in `nets_glm`<sup>5</sup> and we are waiting for this bug to be fixed.

### 5.3. Impact

It is difficult to estimate the impact of our findings, since resting state fMRI entails many different preprocessing choices (e.g. bandpass filtering, global mean regression, nuisance regression) and it is not clear how common different parameters, softwares and pipelines are. For example, it is unknown how common FSLNets (with auto correlation corrected Fisher r-to-z) is compared to REST (Song et al., 2011) and Conn (Whitfield-Gabrieli & Nieto-Castanon, 2012) (both without auto correlation correction). It is also not clear if full correlation is more popular than partial correlation, and how common it is with group analyses where the motion scrubbing substantially differs between two groups. Carp (2012) reviewed methods reporting in the fMRI literature, but did not cover aspects specific to resting state fMRI.

---

<sup>5</sup><http://tiny.cc/n0wwjz>

## Acknowledgements

The authors have no conflict of interest to declare. This study was supported by Swedish research council grant 2017-04889 and CENIIT at Linköping University. Funding was also provided by the ITEA3 / VINNOVA funded project "Intelligence based iMprovement of Personalized treatment And Clinical workflow support" (IMPACT).

## References

- Afyouni, S., Smith, S. M., & Nichols, T. E. (2019). Effective degrees of freedom of the Pearson's correlation coefficient under autocorrelation. *NeuroImage*, *199*, 609–625.
- Arbabshirani, M. R., Damaraju, E., Phlypo, R., Plis, S., Allen, E., Ma, S., Mathalon, D., Preda, A., Vaidya, J. G., Adali, T. et al. (2014). Impact of autocorrelation on functional connectivity. *Neuroimage*, *102*, 294–308.
- Bellec, P., Carbonell, F., Perlberg, V., Lepage, C., Lyttelton, O., Fonov, V., Janke, A., Tohka, J., & Evans, A. (2011). A neuroimaging analysis kit for matlab and octave. In *Proceedings of the 17th International Conference on Functional Mapping of the Human Brain* (pp. 2735–46).
- Bright, M. G., Tench, C. R., & Murphy, K. (2017). Potential pitfalls when denoising resting state fmri data using nuisance regression. *Neuroimage*, *154*, 159–168.
- Carp, J. (2012). The secret lives of experiments: methods reporting in the fMRI literature. *Neuroimage*, *63*, 289–300.
- Chen, G., Saad, Z. S., Nath, A. R., Beauchamp, M. S., & Cox, R. W. (2012). FMRI group analysis combining effect estimates and their variances. *Neuroimage*, *60*, 747–765.
- Craddock, C., Benhajali, Y., Chu, C., Chouinard, F., Evans, A., Jakab, A., Khundrakpam, B. S., Lewis, J. D., Li, Q., Milham, M., Yan, C., & Bellec, P. (2013a). The neuro bureau preprocessing initiative: open sharing of preprocessed neuroimaging data and derivatives. *Neuroinformatics*, .
- Craddock, C., Sikka, S., Cheung, B., Khanuja, R., Ghosh, S. S., Yan, C., Li, Q., Lurie, D., Vogelstein, J., Burns, R. et al. (2013b). Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (c-pac). *Front Neuroinform*, *42*.
- Craddock, R. C., James, G. A., III, P. E. H., Hu, X. P., & Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, *33*, 1914–1928.
- Dansereau, C., Benhajali, Y., Risterucci, C., Pich, E. M., Orban, P., Arnold, D., & Bellec, P. (2017). Statistical power and prediction accuracy in multisite resting-state fMRI connectivity. *Neuroimage*, *149*, 220–232.
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M. et al. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, *19*, 659.
- Eklund, A., Andersson, M., Josephson, C., Johansson, M., & Knutsson, H. (2012). Does parametric fMRI analysis with SPM yield valid results? - An empirical study of 1484 rest datasets. *NeuroImage*, *61*, 565–578.
- Eklund, A., Dufort, P., Villani, M., & LaConte, S. (2014). BROCCOLI: Software for fast fMRI analysis on many-core CPUs and GPUs. *Frontiers in neuroinformatics*, *8*, 24.
- Eklund, A., Knutsson, H., & Nichols, T. E. (2019). Cluster failure revisited: Impact of first level design and physiological noise on cluster false positive rates. *Human brain mapping*, *40*, 2017–2032.
- Eklund, A., Lindquist, M. A., & Villani, M. (2017a). A Bayesian heteroscedastic GLM with application to fMRI data with motion spikes. *NeuroImage*, *155*, 354–369.
- Eklund, A., Nichols, T., Andersson, M., & Knutsson, H. (2015). Empirically investigating the statistical validity of SPM, FSL and AFNI for single subject fMRI analysis. In *International Symposium on Biomedical Imaging (ISBI)* (pp. 1376–1380).
- Eklund, A., Nichols, T., & Knutsson, H. (2016). Cluster failure: why fMRI inferences for spatial extent have inflated false positive rates. *PNAS*, *113*, 7900–7905.
- Eklund, A., Nichols, T., & Knutsson, H. (2017b). Reply to Brown and Behrmann, Cox et al. and Kessler et al.: Data and code sharing is the way forward for fMRI. *PNAS*, *114*, E3374–E3375.
- Fiecas, M., Cribben, I., Bahktiari, R., & Cummine, J. (2017). A variance components model for statistical inference on functional connectivity networks. *NeuroImage*, *149*, 256–266.
- Honari, H., Choe, A. S., Pekar, J. J., & Lindquist, M. A. (2019). Investigating the impact of autocorrelation on time-varying connectivity. *NeuroImage*, *197*, 37–48.
- Lohmann, G., Stelzer, J., Lacosse, E., Kumar, V. J., Mueller, K., Kuehn, E., Grodd, W., & Scheffler, K. (2018). LISA improves statistical analysis for fMRI. *Nature communications*, *9*, 4014.
- Noble, S., Scheinost, D., & Constable, R. T. (2020). Cluster failure or power failure? evaluating sensitivity in cluster-level inference. *NeuroImage*, *209*, 116468.
- Parkes, L., Fulcher, B., Yücel, M., & Fornito, A. (2018). An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *Neuroimage*, *171*, 415–436.
- Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J.-B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*, 115.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, *59*, 2142–2154.
- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*, *84*, 320–341.
- Pruim, R. H., Mennes, M., Buitelaar, J. K., & Beckmann, C. F. (2015). Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *Neuroimage*, *112*, 278–287.
- Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughhead, J., Calkins, M. E., Eickhoff, S. B., Hakonarson, H., Gur, R. C., Gur, R. E. et al. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*, *64*, 240–256.
- Siegel, J. S., Power, J. D., Dubis, J. W., Vogel, A. C., Church, J. A., Schlaggar, B. L., & Petersen, S. E. (2014). Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Human brain mapping*, *35*, 1981–1996.
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., & Woolrich, M. W. (2011). Network modelling methods for FMRI. *Neuroimage*, *54*, 875–891.
- Song, X.-W., Dong, Z.-Y., Long, X.-Y., Li, S.-F., Zuo, X.-N., Zhu, C.-Z., He, Y., Yan, C.-G., & Zang, Y.-F. (2011). REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS one*, *6*, e25031.
- Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain connectivity*, *2*, 125–141.
- Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *Neuroimage*, *21*, 1732–1747.
- Xu, T., Yang, Z., Jiang, L., Xing, X.-X., & Zuo, X.-N. (2015). A connectome computation system for discovery science of brain. *Science Bulletin*, *60*, 86–95.
- Yan, C., & Zang, Y. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Frontiers in Systems Neuroscience*, *4*, 13.
- Yan, C.-G., Cheung, B., Kelly, C., Colcombe, S., Craddock, R. C., Di Martino, A., Li, Q., Zuo, X.-N., Castellanos, F. X., & Milham, M. P. (2013). A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *Neuroimage*, *76*, 183–201.
- Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., & Fischl, B. (2014). Spurious group differences due to head motion in a diffusion MRI study. *Neuroimage*, *88*, 79–90.