Published in final edited form as:

Neuroimage. 2017 January 15; 145(Pt A): 130–132. doi:10.1016/j.neuroimage.2016.10.021.

Corrigendum to "Untangling the relatedness among correlations, Part I: Nonparametric approaches to inter-subject correlation analysis at the group level" [Neuroimage (in press)]

Gang Chen[a,*], **Yong Wook Shin**[b,*], **Paul A. Taylor**[a], **Daniel R. Glen**[a], **Richard C. Reynolds**[a], **Robert B. Israel**[c], and **Robert W. Cox**[a]

[a]Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, USA

[b]University of Ulsan College of Medicine, Department of Psychiatry, Asan Medical Center, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05535 South Korea

[c]Mathematics Department, The University of British Columbia, Canada

1. Introduction

Chen et al. (2016a), hereafter referred to as Part I, is the first in a series of papers describing novel intersubject correlation (ISC) methodologies and analyses. During the analysis of Chen et al. (2016b), referred to as Part II hereafter, it was decided to alter the FMRI preprocessing pipeline originally used in Part I in order to more appropriately handle the parcellation of the task time series as well as implementing a more modern procedure in AFNI (v16.1.16; Cox, 1996). The new sequence of steps was based on the current Example 11 in the afni_proc.py help, which includes, for example, parcellating the time series into movie runs at the start of the analysis, improved alignment to standard space, tissue-based regression using FreeSurfer segmentation (Fischl et al., 2002) and fast ANATICOR (Jo et al., 2010).

Here, we re-analyze the human subject data in Part I using the new FMRI preprocessing, and re-implement each of the ISC methodologies for mutual comparison. We first describe the new set of preprocessing steps, highlighting differences from the description given in Part I (see the second paragraph of the "Performance comparisons with experimental data" section of **Simulations and real experiment results**). The main evaluations and conclusions of Part I are not changed with the new preprocessing steps, but we discuss both the similarities and differences in the observed results.

2. Updated methods

AFNI's afni_proc.py function was implemented for the preprocessing steps, and we note that the succinct command (25 lines) that was used to specify processing steps, to select

*Corresponding authors. gangchen@mail.nih.gov (G. Chen), shaman_korea@mac.com (Y.W. Shin).

desired options, and to generate each subject's full processing pipeline is provided in the Appendix B of Part II, for clarity.

The EPI time series was split into its six sections of movies, each of which was treated as separate scanning run (since the sections were not contiguous; in the earlier analysis, splitting had been performed after removing effects of no interest via regression). In total, there were 406 time points of movie-watching FMRI data. As before, de-spiking and slice timing correction were applied at the start. The subsequent registration steps included motion correction, EPI to anatomical alignment, and standard space registration, all concatenated into single transformations, per time point (to avoid multiple resample operations). The standard space registration was done via anatomical alignment to the Talairach (TT_N27) template using nonlinear registration (providing better alignment than affine) and leaving the EPI data in a 3.5 mm^3 resolution (which was closer to the voxel size at which the data were originally acquired, in lieu of upsampling). Smoothing was performed with a 6 mm isotropic FWHM blur FWHM=6.

Tissue-based regressors were created by first running FreeSurfer's recon-all command to create a parcellation dataset from each subject's anatomical volume, from which lateral ventricle and white matter masks were created and eroded. Regressors were made from the first 3 principal components of the ventricles, as well as applying fast ANATICOR for local white matter regression. Additional regressors included the subject's 6 motion parameters, their derivatives and linear (based on a run length of 142 s) polynomial baseline terms for each of the 6 runs (previously, the scan had been treated as a single, continuous time series), for a total of 28 regressors. Time points were censored from the regression model whenever the per-time point motion (Euclidean norm of the motion derivatives) exceeded 0.3 mm or when at least 10% of the brain voxels were seen as outliers from the trend. Finally, as before, ISC was computed over the final time series (here, having 406 time points) using 3 dTcorrelate in AFNI. The same group analyses as described in Part I were performed on the N=48×47/2=1128 ISC values per voxel.

All the bootstrap and permutation approaches (SWB, EWB, SWP and EWP) as well as t-tests for ISC were implemented on the newly processed data, as described in Part I. The ISC Toolbox and 3dttest++ were also re-run, and all statistical analyses and comparisons for the experimental data were performed in the same manner as in Part I. The formats of the figures shown here are directly copied from those shown in Figs. 3 and 4 of Part I, utilizing the same format, arrangement, statistical thresholds, etc.

3. Results

The ISC results are provided here in Figs. 1A and 2 (mirroring Fig. 3 and 4, respectively, in Part I), where the Pearson correlation values are shown in voxels whose significances were thresholded at $p < 0.001$ (for one group; top row) and $p < 0.05$ (for two groups; bottom row). Fig. 1B contains counts of voxels within the whole brain that were above threshold for each case.

The results are qualitatively similar to those described in Part I: relative to the other one-group comparisons, SWB provides much more specific clusters. This is observed in both the axial slice images and in the suprathreshold voxel counts. Interestingly, compared to Part I results, the re-processed data provide thresholded clusters that are much more localized to GM, and in particular to regions of networks expected to be associated with the movie-watching task (i.e., visual and audio). While the volumes of statistical significance for previously existing methods are less "overwhelming" than those in Part I, it can be observed in the voxel counts that localization for most methods is still extremely difficult, even for small p-values. The importance of accounting for the correlation structure $\boldsymbol{P}^{(n)}$ using "subject-wise" randomization is still apparent (e.g., see SWB in the one group tests).

4. Conclusions

It is not possible to assign the contributions of individual processing step changes to differences in the output data. However, we would strongly recommend the choice of the several options implemented here (nonlinear alignment to template space, individual regressors for each movie clip, using ANATICOR regressors) for each of the relevant steps in FMRI processing.

Importantly, the main results and their interpretation from the original Part I remain the same. The relative specificity of SWB compared to other methods was expected from the comparisons of the approaches using the simulations, and this approach remains the recommended one for ISC processing of one group (for two groups, SWP had been shown to provide the best agreement with nominal false positive rates for two group comparisons). We note that Part II provides a linear mixed effects (LME) strategy for ISC analysis, which is compared to the subject-wise methods of this study; in that work features of the derivation and output of LME modeling are shown to provide additional information for the interpretation of even the nonparametric results, shown in here and in Part I.

References

Chen G, Shin YW, Taylor PA, Glen DR, Reynolds RC, Israel RB, Cox RW. Untangling the relatedness among correlations, part I: nonparametric approaches to inter-subject correlation analysis at the group level. Neuroimage. 2016a [in press].

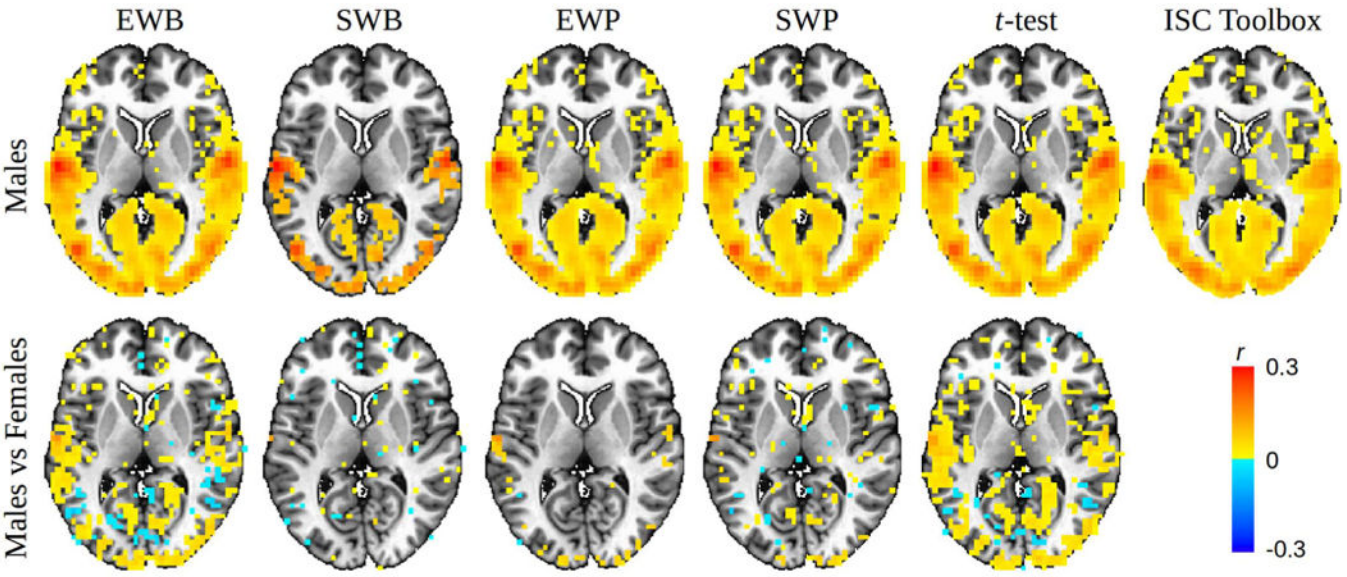
Chen GC, Taylor PA, Shin YW, Reynolds RC, Cox RW. Untangling the relatedness among correlations, part II: inter-subject correlation group analysis through linear mixed-effects modeling. Neuroimage. 2016b [under review].

Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res. 1996; 29:162–173. <http://afni.nimh.nih.gov/afni>. [PubMed: 8812068]

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002; 33(3):341–355. [PubMed: 11832223]

Jo HJ, Saad ZS, Simmons WK, Milbury LA, Cox RW. Mapping sources of correlation in resting state FMRI, with artifact detection and removal. Neuroimage. 2010; 52(2):571–582. [PubMed: 20420926]

A) Visual comparisons of one slice.



B) Numbers of suprathreshold voxels at various significance levels.

Significance Level	One Group: Males						Group Comparison: Males vs Females				
	EWB	SWB	EWP	SWP	*t*-test	ISC Toolbox	EWB	SWB	EWP	SWP	*t*-test
0.05	652,172	327,608	772,951	760,260	711,725	827,729	218,748	57,024	31,556	105,087	240,872
0.01	535,380	195,553	647,455	623,831	576,840	669,752	119,193	11,362	17,107	28,298	126,267
0.005	500,223	160,781	607,882	578,770	536,195	*	97,069	4,888	14,492	17,064	101,271
0.001	435,353	105,215	534,737	489,375	462,879	539,946	64,741	1,243	10,590	6,260	65,342
0.0005	419,789	93,339	517,673	466,823	438,140	*	59,168	858	9,947	5,188	54,537

Fig. 1.

Performance comparisons with an experimental dataset. (A) Axial views (Z=8 mm; radiological convention: left is right) of ISC group results (thresholded by p-values, below) of an experimental dataset are illustrated for the five methods as well as the approach implemented in the ISC Toolbox. The colors code for the magnitude of correlation coefficients. For both one- and two-sample tests, the results are consistent with their FPR controllability. Specifically, for the male group (n= 24, upper panel, two-tailed significance level p=0.001) all the other five methods are more liberal than SWB. Although not visually obvious in the color coding, the group ISC estimates through averaging across subjects in the ISC Toolbox tend to be biased relative to the medians from the other four nonparametric methods. For two-group comparison (n_1=n_2=24, lower panel, two-tailed significance level p=0.05; group comparison testing currently not available in the ISC Toolbox), EWP was much more liberal relative to SWP for some regions, while being over-conservative for others; EWB was too liberal, SWB tended to be slightly more conservative, and both rendered noisier results; and Student's t-test was the worst. The SWP performance for the other two indirect contrasts ($\boldsymbol{R}_{11}$ vs $\boldsymbol{R}_{21}$ and $\boldsymbol{R}_{22}$ vs $\boldsymbol{R}_{21}$) are shown in Fig. 2. We note that 1) multiple testing correction was not performed so that voxel-wise comparisons among the methods could be directly visualized; and 2) except for the ISC Toolbox, which adopts direct averaging across ISC values, all other methods rendered virtually the same group estimate for ISC, but differed in significance detection (i.e., the color at each voxel is roughly the

same across the first five testing methods if the significance survives the corresponding threshold). The performance comparisons among the five methods with the BGC component $\boldsymbol{R}_{21}$ (not shown here) are similar to the one group scenario for males (first row), $\boldsymbol{R}_{11}$. (B) List of voxels (resolution: in 1×1×1 mm^3) that pass five significance levels. The default setting in ISC Toolbox does not provide thresholding at voxel-wise significance levels of 0.005 and 0.0005 (marked with * in the table). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

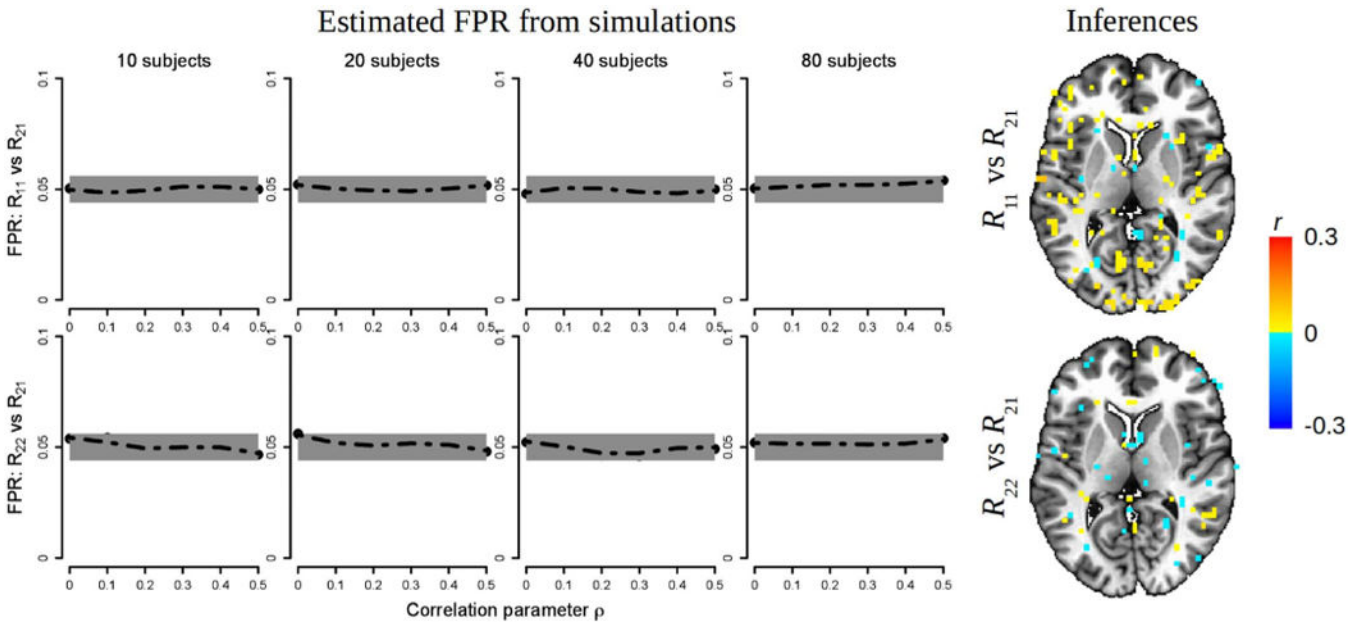


Fig. 2.

Performances of SWP on the comparisons of within-group versus between-group ISCs (i.e., indirect comparisons) are shown in terms of FPR controllability (left) and inferences (right) with an experimental dataset. The simulation parameters and experimental dataset are the same as in Fig. 2 in Part I and Fig. 1 here, respectively, but the scale for FPR (left) is different from Fig. 2 in Part I. The gray line of FPR=0.05 indicates the 95% confidence band of the target (or nominal) value (with a band width of 0.012 for each simulation with 5000 realizations).