*Under consideration for publication in Network Science* 

# *A Paradigm for Longitudinal Complex Network Analysis over Patient Cohorts in Neuroscience : Supplementary Appendices*

Heather Shappell *
Department of Biostatistics, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD
Yorghos Tripodis
Department of Biostatistics, Boston University, Boston, MA
Ronald J. Killiany
Department of Anatomy and Neurobiology, Boston University, Boston, MA
Eric D. Kolaczyk
Department of Mathematics and Statistics, Boston University, Boston, MA
and for the Alzheimers Disease Neuroimaging Initiative **

## A Additional data and pre-processing information

All of the rsfMRI scans were visually inspected to ensure that they were free from any artifacts (i.e. pencil beam artifact). No scans were excluded due to artifact. FMRI Expert Analysis Tool (FEAT; Oxford, UK; v6.0 http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL) was used for this preprocessing (Smith *et al.*, 2009). Single-session independent component analysis (ICA) was run on the rsfMRI data with MELODIC v3.14. The first ten volumes of each scan were discarded. Preprocessing included highpass filtering (cutoff=100s), motion correction (MCFLIRT), and spatial smoothing (FWHM=5mm). Time courses were variance-normalized, and automatic dimensionality estimation was performed, resulting in each participants data being separated into a different number of components. The preprocessed data output from MELODIC were entered into FSLs FIX v1.06 and processed using the Standard.RData trained-weights file, a threshold of 20, and additional motion cleanup. FIX was used to remove noise from the data; particularly noise resulting from motion, susceptibility, cardiac pulsations, white matter, and cerebrospinal fluid. Edits were made as needed and the edited components were regressed out of the rsfMRI data.

Each rsfMRI sequence was registered to the brain image extracted from the structural scan by Freesurfer and the resulting rsfMRI sequence was labeled using the generated Freesurfer ROIs. A mean time series for each ROI was calculated by averaging all fMRI voxel values within each ROI over time, resulting in 140 time points calculated for each 7 minute resting state session. The final list of ROIs used in our analyses are shown on the following page.

- Left-Hippocampus
- ctx-lh-bankssts
- ctx-lh-caudalanteriorcingulate
- ctx-lh-caudalmiddlefrontal
- ctx-lh-cuneus
- ctx-lh-entorhinal
- ctx-lh-fusiform
- ctx-lh-inferiorparietal
- ctx-lh-inferiortemporal
- ctx-lh-insula
- ctx-lh-isthmuscingulate
- ctx-lh-lateraloccipital
- ctx-lh-lingual
- ctx-lh-middletemporal
- ctx-lh-paracentral
- ctx-lh-parahippocampal
- ctx-lh-parsopercularis
- ctx-lh-parsorbitalis
- ctx-lh-parstriangularis
- ctx-lh-pericalcarine
- ctx-lh-postcentral
- ctx-lh-posteriorcingulate
- ctx-lh-precentral
- ctx-lh-precuneus
- ctx-lh-rostralanteriorcingulate
- ctx-lh-rostralmiddlefrontal
- ctx-lh-superiorfrontal
- ctx-lh-superiorparietal
- ctx-lh-superiortemporal
- ctx-lh-supramarginal
- ctx-lh-temporalpole
- ctx-lh-transversetemporal
- Right-Hippocampus
- ctx-rh-bankssts
- ctx-rh-caudalanteriorcingulate
- ctx-rh-caudalmiddlefrontal
- ctx-rh-cuneus
- ctx-rh-entorhinal
- ctx-rh-fusiform
- ctx-rh-inferiorparietal
- ctx-rh-inferiortemporal
- ctx-rh-insula
- ctx-rh-isthmuscingulate
- ctx-rh-lateraloccipital
- ctx-rh-lingual
- ctx-rh-middletemporal
- ctx-rh-paracentral
- ctx-rh-parahippocampal
- ctx-rh-parsopercularis
- ctx-rh-parsorbitalis
- ctx-rh-parstriangularis
- ctx-rh-pericalcarine
- ctx-rh-postcentral
- ctx-rh-posteriorcingulate
- ctx-rh-precentral
- ctx-rh-precuneus
- ctx-rh-rostralanteriorcingulate
- ctx-rh-rostralmiddlefrontal
- ctx-rh-superiorfrontal
- ctx-rh-superiorparietal
- ctx-rh-superiortemporal
- ctx-rh-supramarginal
- ctx-rh-temporalpole
- ctx-rh-transversetemporal

## B Sample Size Calculation

As an illustration of a sample size calculation in the context of our study, we have gone ahead and performed a rough power calculation. We would need approximately 190 subjects for the case/control parameter estimate associated with the DMN effect outcome to be significant at the $\alpha = 0.05/6 = 0.0083$ level (i.e. the Bonferroni adjusted $\alpha$ level assuming our main interest is in the case/control estimates). This calculation is assuming an effect size of 0.16 (based off of our primary analysis), 80% power, 3 observations per subject, and a conservative correlation of 0 among repeated measures. If we increase the correlation among repeated measures to be 0.4 (which is much larger than what it estimated from our current data), the sample size drops to be 116. G* Power 3 was used for these calculations (Faul *et al.*, 2007).

## C Secondary Analysis Results

In this analysis, we no longer define the caudal anterior cingulate as a DMN ROI, and we add the superior frontal as a DMN ROI. We keep the remainder of the DMN list (as defined in section 3.3.2) the same. The results of this analysis are displayed below in Table C1. We again find a significant interaction between disease status and time period for the three-cycles and between effects (plots are shown in Figures D1 and D2). We are no longer seeing a significant interaction between disease status and time decrease for the bilateral effect, so the interaction term has been removed from that model. Probably the most notable change in this analysis is that both the DMN and same lobe effects now have a significant interaction between disease status and time period at the $\alpha = .10$ significance level. The mean values of each over time (for both the primary and secondary analysis) are shown in Figures D4 and D5. This gives a clearer picture of what is happening and one can see that the results actually do not end up looking so different from the primary analysis. The same lobe effect decreases over time for both cases and controls in both analyses. The exception to this is sometime period 4, but the sample size at period 4 is much smaller than the other periods for the AD group (n=3). The controls have larger effects over time compared to the cases in both. For the DMN effect, time does not appear to play much of a role (i.e. the change in the effect over time is minimal) and controls do appear to consistently have larger effect values than cases. The trends aren't quite as parallel in the secondary analysis as they were in the first, which is why we are seeing a potential interaction in the numerical results. However, as can be seen from the plots, the interaction is more quantitative than qualitative. If we remove the interaction, we still see a difference in the DMN and same lobe effects in cases versus controls, but the p-values aren't quite as small (p-values = .0251 and 0.106, respectively).

Table C 1: Meta-Analysis Results for the Secondary Analysis.

| Effect | Intercept | Case[a] | SE[b] | T[c] | P[d] | Period Est[e] | Period P[f] | Interact Est[g] | Interact P[h] |
|---|---|---|---|---|---|---|---|---|---|
| Three-Cycles | 0.109 | 0.021 | 0.010 | 2.19 | 0.031 | 0.0005 | 0.831 | -0.009 | 0.012 |
| Distance 2 | -0.144 | 0.005 | 0.009 | 0.55 | 0.585 | | | | |
| Between | -0.135 | 0.003 | 0.011 | 2.99 | 0.004 | 0.008 | 0.004 | -0.016 | $<$0.001 |
| Bilateral | 1.050 | 0.076 | 0.116 | 0.66 | 0.513 | | | | |
| Same Lobe | 0.383 | 0.020 | 0.045 | 0.45 | 0.651 | -0.023 | 0.022 | -0.032 | 0.060 |
| DMN | 0.209 | 0.079 | 0.113 | 0.70 | 0.488 | 0.032 | 0.227 | -0.103 | 0.027 |

[a] Controls are the reference group.
[b] Standard error corresponding to the case/control parameter estimate.
[c] T-ratio corresponding to the case/control parameter estimate. Calculated prior to rounding.
[d] P-value corresponding to the case/control parameter estimate.
[e] Parameter estimate associated with the period number.
[f] P-value corresponding to period number parameter estimate.
[g] Case$\times$Period interaction parameter estimate.
[h] Case$\times$Period interaction p-value.

## D Additional Figures

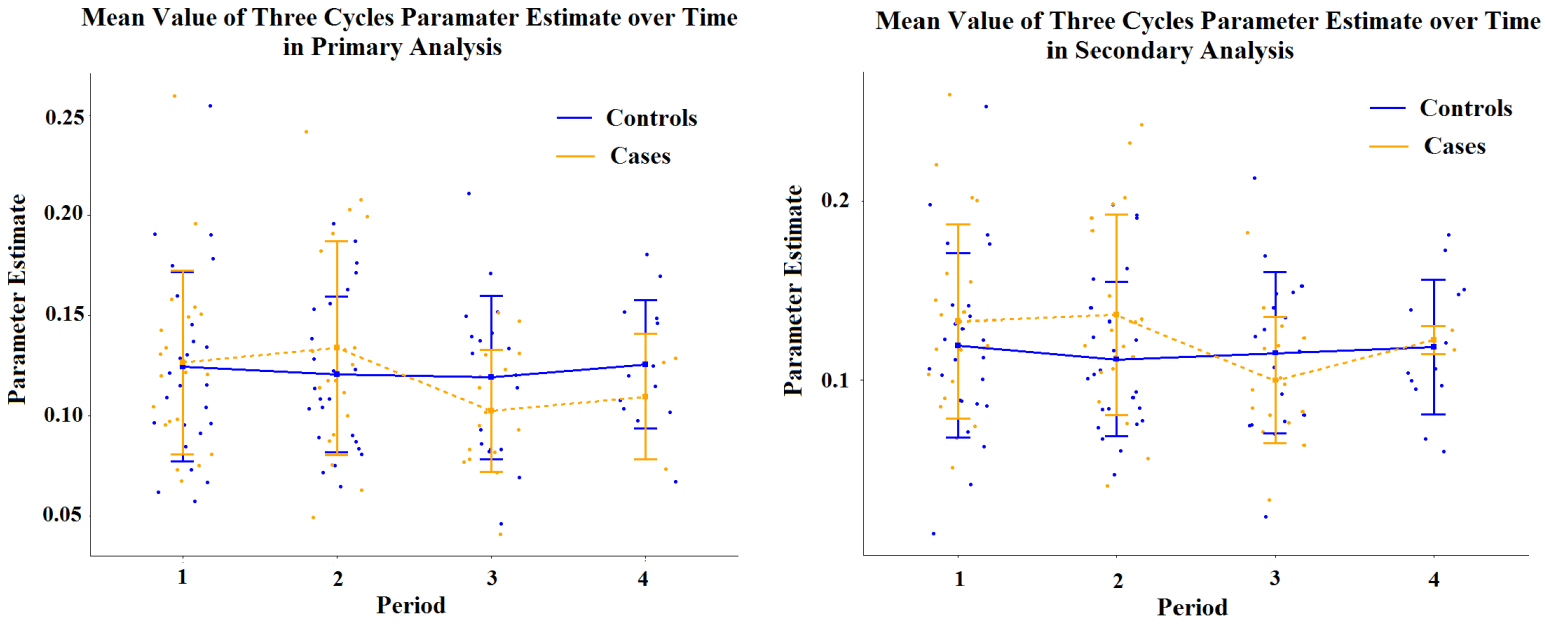


Fig. D 1: Mean value of three cycles effect at each time period for the primary and secondary analyses.

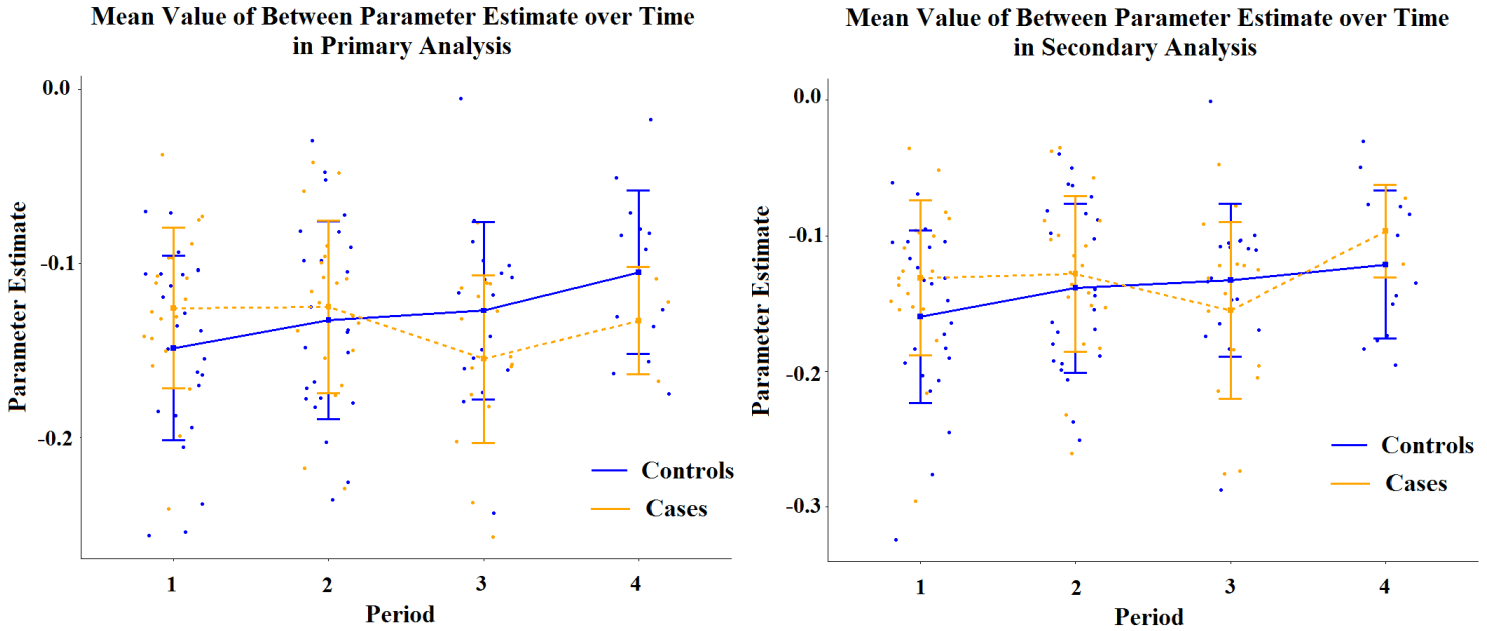


Fig. D 2: Mean value of between effect at each time period for the primary and secondary analyses.

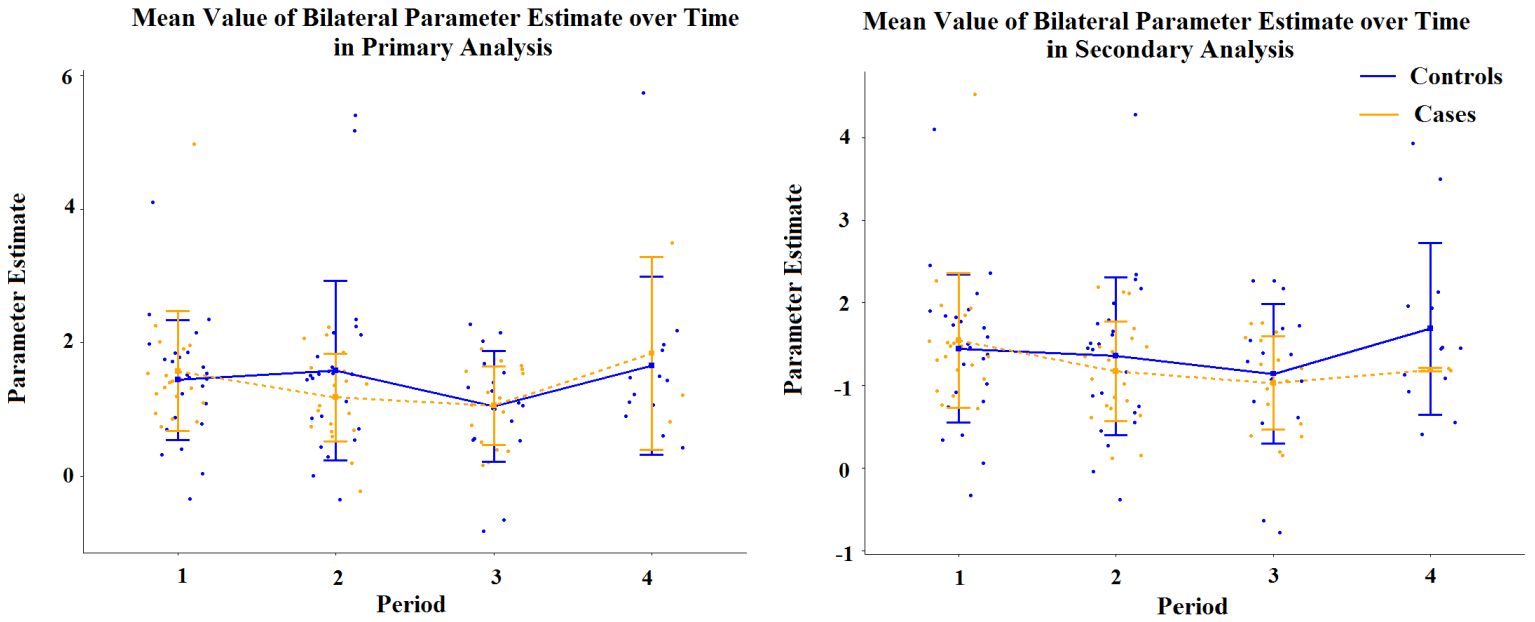


Fig. D 3: Mean value of bilateral effect at each time period for the primary and secondary analyses.

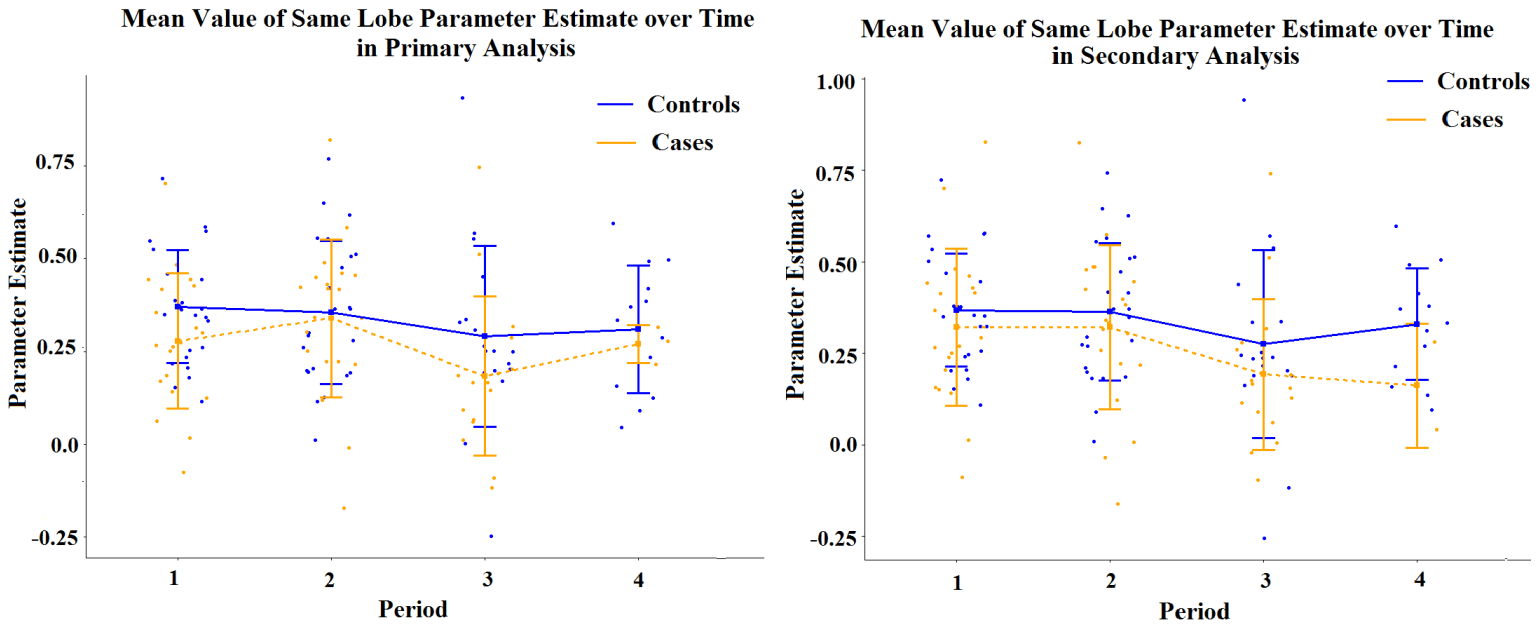


Fig. D 4: Mean value of same lobe effect at each time period for the primary and secondary analyses.

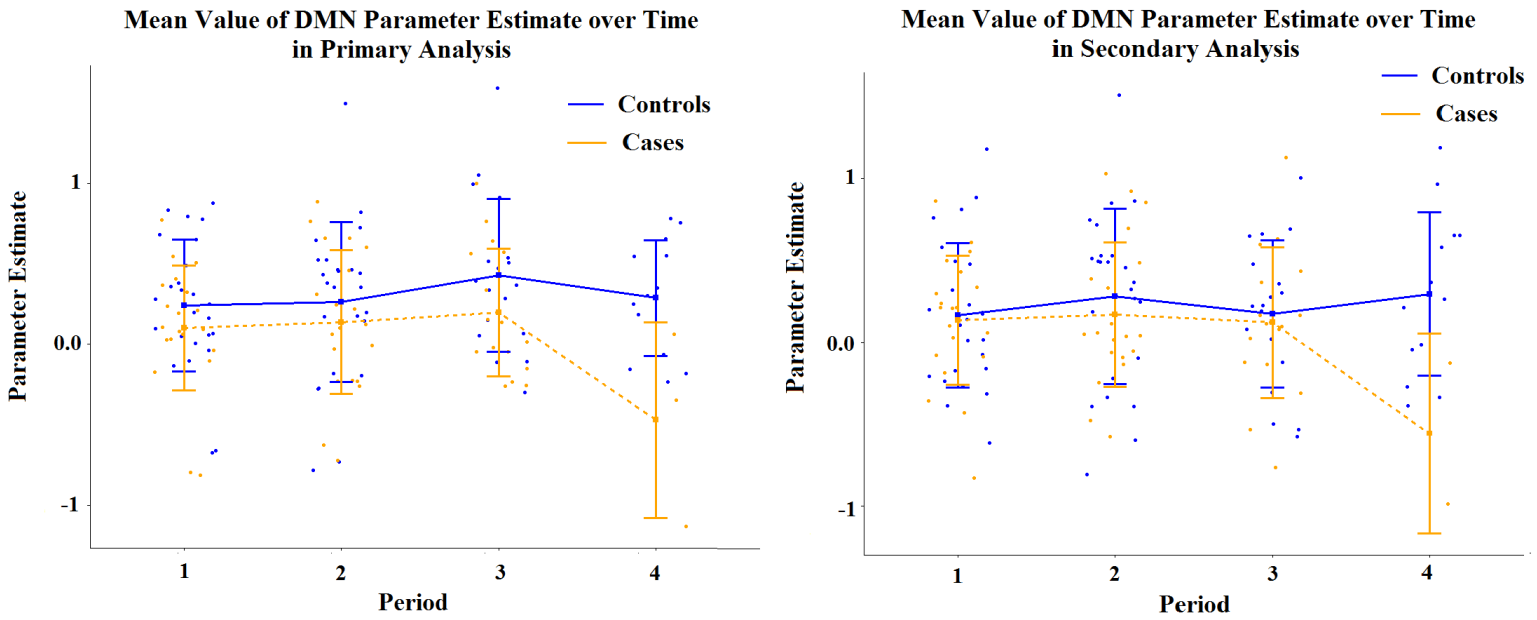


Fig. D 5: Mean value of DMN effect at each time period for the primary and secondary analyses.

## References

Faul, Franz, Erdfelder, Edgar, Lang, Albert-Georg, & Buchner, Axel. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, **39**(2), 175–191.

Smith, Stephen M, Fox, Peter T, Miller, Karla L, Glahn, David C, Fox, P Mickle, Mackay, Clare E, Filippini, Nicola, Watkins, Kate E, Toro, Roberto, Laird, Angela R, *et al.* . (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the national academy of sciences*, **106**(31), 13040–13045.