Coupled Intrinsic Connectivity: A Principled Method for Exploratory Analysis of Paired Data

Dustin Scheinost, Xilin Shen, Emily Finn, Rajita Sinha,
R. Todd Constable, and Xenophon Papademetris

Departments of Biomedical Engineering, Diagnostic Radiology,
Interdepartmental Neurosceince Program, Psychiatry, and Nuerosurgery
Yale University, New Haven, CT

Abstract. We present a novel voxel-based connectivity approach for paired functional magnetic resonance imaging (fMRI) data collected under two different conditions labeled the Coupled Intrinsic Connectivity Distribution (coupled-ICD). Our proposed method jointly models both conditions to incorporate additional spatial information into the connectivity metric. When presented with paired data, conventional voxel-based methods analyze each condition separately. However, nonlinearities introduced during processing can cause this approach to underestimate differences between conditions. We show that commonly used methods can underestimate functional changes and evaluate our coupled-ICD solution using a study comparing cocaine-dependent subjects and healthy controls. Our approach detected differences between paired conditions in similar brain regions as the conventional approaches while revealing additional changes. Follow-up seed-based analysis confirmed, via cross validation, connectivity differences between conditions in regions detected by coupled-ICD that were undetected using conventional methods. This approach of jointly analyzing paired connectivity data provides a new and important tool with many clinically relevant applications.

Keywords: Functional Connectivity, Resting-state, fMRI, Cocaine

1 Introduction

Functional connectivity holds promise as a clinical tool to detect abnormal brain organization in clinical populations. The most common approaches rely on regions of interests (ROIs) or "seeds" to characterize connectivity; however, seed-based approaches can only examine connectivity in reference to the seed region. Choosing which seeds to examine is often a difficult question as the wrong choice in seed regions could occlude important patterns of connectivity.

Voxel-based metrics can be used as a data-driven way to define seed regions for further analysis [1–3]. However, voxel-based approaches and seed approaches can often produce seemingly conflicting results. For example, voxel-based results may suggest an increase in connectivity for a region while follow-up seed analysis with the region may show decreases in connectivity to the region. This discrepancy arises because each approach is fundamentally different. Voxel-based

metrics essentially work as compression mechanisms, reducing all information about the connections to a voxel into a few summary parameters used for group comparisons. This compression of information is nessessary as connectivity matrix genernated from voxel-based approaches (typically 20,000x20,000 matrix at fMRI resolutions) are difficult to interpret and are problematic for statistical inferences. In contrast, seed-based approaches directly compare correlations between regions at the group level rather then these summary parameters.

For the special case of paired data such as pre- and post-treatment, the standard approach with voxel-based metrics is to compure the summary parameters separately for each condition and then perform statistical analysis to compare the two. We observe that this approach is suboptimal as the compression into a summary parameters is performed twice (once for each condition). Thus, with this approach, comparisons are made on how these summary parameters change rather then how correlations between regions change. Further, non-linearities introduced during processing – such as only examining the positive correlations [1, 4] – guarantee that the difference in the summary parameter is not the same as the summary parameter of a difference. Hence, information about how each correlation changes due to the treatment is also lost with current approaches.

In this work, we propose a method where within-subject differences across conditions are first computed and then a single summary measure can be calculated for these differences. We label our approach the coupled Intrinsic Connectivity Distribution (coupled-ICD) as it extends the recently developed Intrinsic Connectivity Distribution (ICD) method [9]. Unlike other voxel-based metrics, coupled-ICD mimics seed-based approaches by directly comparing correlations between each condition and, then, summarizing these changes into summary parameters for group comparisons. Thus, coupled-ICD should produce regions more suitable for seed-based connectivity. To assess our coupled-ICD measure, we used a data set of cocaine-dependent subjects and healthy controls scanned while presented with relaxing and drug-related imagery cues. We show that our coupled-ICD has higher sensitivity than conventional approaches for detecting differences between conditions. Finally, using cross-validation on separate, independent sub-sample of our data, we show regions detected by coupled-ICD are predictive of seed-based difference in connectivity.

2 Theory

Voxel-based measures of functional connectivity [4, 1, 9] aim to reduce large amounts of information to a voxel into a much smaller set of summary parameters. Typically, this compression is formulated based on graph theory [8] where the brain is treated as a graph or network and each voxel represents a node in this graph. These nodes (or voxels) are connected to each other by edges based on the similarity of their timecourses.

Measures of node centrality such as the measure degree are the primary metrics used for compression. For any voxel x, these measures can be calculated from the distribution of connection strength, $f(x, r)$, where x is the current

voxel, and r is a correlation or any other measure of timecourse similarity. First, $f(x, r)$ is estimated by computing the histogram of the correlations r for the timecourse at voxel x to the timecourse at every other voxel in the brain. Degree can be estimated as the integral of this distribution from any threshold τ to 1, or $\int_{\tau}^{1} f(x, r)dr$. Weighted degree measures such as weighted Global Brain Connectivity (wGBC) [1] can be estimated as the mean of this distribution. In contrast, ICD models the corresponding survival function to $f(x, r)$. Each point on the survival function is simply degree evaluated at that particular threshold τ and, thus, ICD parameterizes how degree for a voxel changes as the threshold used to determine if two voxels are connected is increased. Previously, it was shown that a stretch exponential decay with unknown variance parameter α and shape parameter β was sufficient to model this survival function. Modeling the survival function with a stretch exponential is equivalent to modeling the underlying distribution as a Weibull distribution: $f(x, r, \alpha, \beta) = \frac{\beta}{\alpha}(\frac{r}{\alpha})^{\beta-1} \exp(-(\frac{r}{\alpha})^{\beta})$.

The presented approach, coupled-ICD, extends conventional voxel-based connectivity in a critical way as the graph summarized by coupled-ICD is a graph defined by differences in correlations and not simply the correlations. As such, coupled-ICD takes advantage of the paired nature of the data by explicitly comparing the same edge under two conditions. Typically, graphs at the voxel level become difficult to analyze due to memory constraints and multiple comparison issues with groups of 10 or more subjects. However, for the special case of paired scans, only two graphs need to be analyzed simultaneously. Paired scans reduce the complexity of this problem.

Directly comparing each correlation (like seed-based appoarches) allows additional spatial information about changes of corresponding edges to be incorporated into a summary parameter of connectivity for a single subject. Specifically, computing the differences between the weights of corresponding edges of a graph and then summarizing the differences takes into account the topological (spatial) structure of the graphs. With this approach, information about how each edge has changed due to condition or time can be incorporated into the summary parameter for the subject. This information is lost with current approaches. Due to non-linearities in the calculation of these summary parameters, the difference between the summary parameters (degree, wGBC, or ICD) of two graphs is not the same as the summary parameter of the difference between the graphs. For example, given the ambiguity of negative correlations, many current approaches examine only the positive correlations [4, 9, 1].

Given a set of paired data, coupled-ICD can be computed by repeatedly calculating conventional seed connectivity maps treating each voxel as a seed, and summarizing the difference between the seed maps for each condition (Fig. 1). First, for any voxel x, the correlation between the timecourse at voxel x to the timecourse at every other voxel in the gray matter is calculated for each condition in the paired data. These correlation maps are then subtracted from one another. Coupled-ICD then summarizes this map of differences in the same way that ICD (or degree) summarizes a map of connections to a voxel. First, for each voxel, a distribution of these differences is estimated with a histogram. Second, this dis-

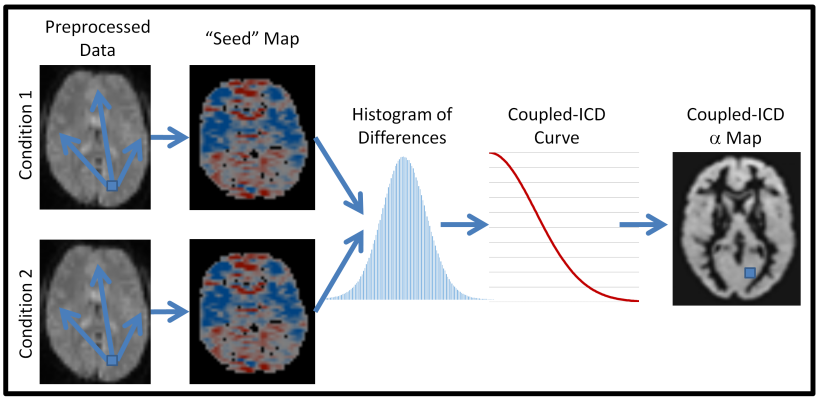

Fig. 1. Flow chart describing coupled-ICD. For paired data, coupled-ICD jointly analyzes both conditions and then creates a summary of the difference in connectivity between conditions for each voxel. First, a seed connectivity map is created for a voxel (shown as the blue square through the flow chart) in each condition. The resulting survival function of the distribution of the difference (labeled coupled-ICD curve) is calculated and modeled with a stretched exponential. This process is repeated for each voxel in the gray matter. The final output is an image where each voxel represents a summary of the difference between two seed maps using that voxel as the seed region.

tribution is modeled as a Wiebull distribution which corresponds to modeling the survival function of the histogram as a stretch exponential. Group comparisons can be performed by comparing the parameters with standard methods.

Coupled-ICD can be used to model increases in connectivity, decreases in connectivity, or the magnitude of the changes in connectivity. For simplicity, we focus only on modeling magnitude of the changes in connectivity noting that the modeling and the interpretation of the parameters is similar for the other cases. When modeling the magnitude of changes, a larger α parameter indicates a larger variance in the distribution and that a larger number of connections exhibit a strong change in correlations between the two conditions.

3 Functional Connectivity Estimation

Subjects: The data set consisted of 28 cocaine-dependent (CD) subjects and 38 healthy control (HC) subjects aimed at examining influence of cue state and diagnostic group on brain activity. Subjects performed four fMRI scans while listening to imagery scripts of either neutral relaxing cues or drug related cues (two scans of each). Complete details can be found elsewhere [7, 10].
Preprocessing: Images were slice-time corrected using sinc interpolation and motion corrected using SPM5. All further analysis was performed using in-house software. Several covariates of no interest were regressed from the data including linear and quadratic drift, six rigid-body motion parameters, mean cerebral-spinal fluid (CSF) signal, mean white matter signal and mean global signal.

Finally, the data were temporally smoothed with a zero mean unit variance Gaussian filter (cutoff frequency=0.12Hz). A gray matter mask was applied to the data so that only voxels in the gray matter were used in the calculation.

ICD and Degree: The timecourse for voxel x was correlated with the timecourse for every other voxel in the gray matter. As the removal of the global mean makes the signs of the correlation ambiguous [4], only the positive correlation was used in analysis. For each voxel, a distribution of connection strength was estimated for the positive correlation coefficients using a 100 bin histogram. ICD was used to model this distribution. First, the histogram was converted to the corresponding survival function and this survival function was modeled with a stretched exponential. This results in two summary parameters for each voxel reflecting that voxels connectivity to the rest of the brain; the α parameter was used in the group comparisons. As this survival function describes how the network theory measure degree changes with connection threshold, degree at any threshold can be estimated as a single point on the survival function. Degree was estimated with a connection threshold of $r = 0.25$. ICD or degree maps for each paired condition were then subtracted from each other resulting in a single map per subject describing the difference in connectivity between conditions.

Coupled-ICD: Similar to the ICD and degree estimations, the timecourse for voxel x was correlated with the timecourse for every other voxel in the gray matter. As coupled-ICD operates on paired data, this process was performed on both conditions resulting in two seed connectivity maps with voxel x as the seed. These maps are then subtracted. A distribution of the differences in connection strength was estimated for the absolute value of the differences using a 200 bin histogram. A larger number of bins was used to keep the bin width the same as the ICD analysis while accommodating the wider range of possible values (difference in correlations has a range of $[-2, 2]$ while correlation has a range of $[-1, 1]$). We chose to model the absolute value of the differences to highlight regions of the brain that show large differences between two conditions. A similar procedure can be used to analyze the increases or decreases between conditions. As described above, this histogram was converted to the corresponding survival function and this survival curve was modeled with a stretched exponential.

Seed Connectivity: A follow-up seed-based analysis (similar to [4, 2, 3]) was performed on a sample region showing large differences as detected using coupled-ICD. The voxel-based analysis and follow-up seed-based analysis were run on independant data by spliting the data into two groups by subjects. Fourteen CD subjects and 19 HC subjects were randomly chosen for voxel-based analysis. The remaining subjects were used for seed-based analysis. Splitting the data into two groups allows the seed connectivity results to act as a cross-validation of the coupled-ICD results. A seed was placed in putamen based on voxels showing significant differences ($p < 0.05$, corrected) between HC and CD subjects. The timecourse of the reference region in a given subject was then computed as the average timecourse across all voxels in the reference region. This timecourse was correlated with the timecourse for every other voxel in the gray matter to cre-

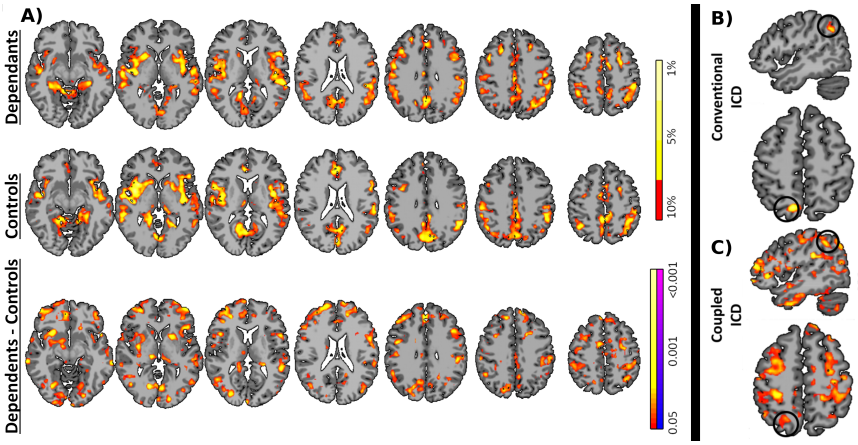

Fig. 2. Evaluation of coupled-ICD. A) Coupled-ICD detects widespread differences in connectivity while subjects are experiencing either relaxing or drug-related imagery for both (*top*) the cocaine-dependent (CD) subjects. and (*middle*) healthy controls (HC). (*Bottom*) Group level comparisons (CD vs HC) revealed that the groups significantly ($p < 0.05$, corrected) differ in response to the imagery conditions. **B)** Conventional ICD analysis detected significant group differences ($p < 0.05$ corrected) in the parietal and occipital lobes. **C)** Coupled-ICD detected significant ($p < 0.05$, corrected) difference in these areas as well as several other areas. This result highlights the additional information that can be captured by jointly analyzing paired conditions.

ate a map of r-values, reflecting Seed-to-whole-brain connectivity. These r-values were transformed to z-values using Fisher's transform.

Group Analysis: To facilitate group statistics, all single subject results were spatially smoothed with a $6mm$ Gaussian filter and non-linearly warped to common space using Bioimage Suite [5]. Between group differences were calculated using two-sample t-test with significance assessed at $p < 0.05$. AFNI's Alphasim was used for multiple comparison correction.

4 Results

Coupled-ICD detected widespread significant differences due to condition between the two groups. The coupled-ICD maps for each group and the between-group comparisons are shown in Fig. 2. Both groups showed large differences between the drug-related and relaxing imagery in the posterior cingulate cortex, bilateral angular gyrus, bilateral insular cortex, bilateral putamen, medial prefrontal/anterior cingulate cortex, and visual processing areas. The CD group showed significantly greater differences between conditions in many of these regions and additionally in the prefrontal lobe.

The between-group differences detected by coupled-ICD were compared with between-group differences detected by the conventional ICD and degree approaches. Conventional ICD detected two clusters that satisfied our criteria for

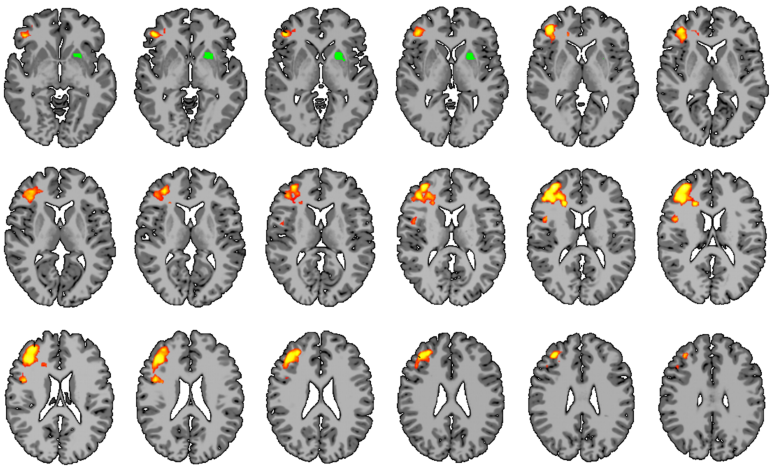

Fig. 3. Seed validation. A follow-up, seed-based connectivity analysis was performed on independant data using a region in the left insula detected by coupled-ICD but not by conventional ICD and degree analysis. Several areas of significant differences were detected ($p < 0.05$, corrected). As the seed analysis was performed on independant data, the seed-based results provide evidence via cross-validation that coupled-ICD is detecting an effect not detected by conventional analysis

significant differences-in the parietal and occipital lobes-while the degree approach did not detect any significant differences. Group differences for both coupled-ICD and ICD are shown in Fig. 2. Coupled-ICD identifies both clusters detected by conventional ICD along with additional widespread changes.

To explore the differences detected by coupled-ICD, a follow-up seed-based analysis was performed using a seed defined in the left putamen where significant between-group differences were found using coupled-ICD. The left putamen was chosen as a seed due to the substantial literature (see [7] for example) implicating this region in addiction. Significant ($p < 0.05$, corrected) interactions between group and condition were observed in right frontal lobe (Fig. 3).

5 Discussion

We present a principled method for exploratory analysis of paired conditions to detect regions that differ significantly in their connectivity patterns between conditions. We show that our coupled-ICD approach is superior in detecting group differences in connecitivity due to paired conditions. We show that coupled-ICD is a viable solution as a data-driven way to pick seeds for further analysis. Standard seed-based analysis, performed on data independent from the coupled-ICD results, showed that the regions detected by coupled-ICD exhibit significant differences in seed connectivity. While similar to conventional voxel-based metrics of connectivity, the present approach extends connectivity in a critical

way, coupled-ICD summaries differences in correlations, rather compute difference in summaries of correlation. This extension allows coupled-ICD to mimic seed-based approaches and gain additional information for group comparisons.

Numerous clinical applications could benefit from measuring changes in the functional organization of the brain at the voxel level for paired data, yet the translational technology for detecting changes in connectivity remains elusive. Coupled-ICD represents a principled method for exploratory analysis of paired conditions to detect regions that differ significantly in their connectivity patterns between conditions. Thus, coupled-ICD could potentially fill this important void not currently covered by conventional approaches.

Acknowledgments This work was funded in part by NIH R01 NS052344, R01 EB00966, R03 EB012969.

References

1. Cole, M.W., Yarkoni, T., Repovs, G., Anticevic, A., Braver, T.S. (2012) Global connectivity of prefrontal cortex predicts cognitive control and intelligence. The Journal of Neuroscience, 32:8988-99.
2. Constable, R.T., Vohr, B.R., Scheinost, D., Benjamin, J.R., Fulbright, R.K., Lacadie, C., Schneider, K.C., Katz, K.H., Zhang, H., Papademetris, X., Ment, L.R. (2012) A left cerebellar pathway mediates language in prematurely-born young adults. Neuroimage, 64C:371-378.
3. Hampson, M., Tokoglu, F., Shen, X., Scheinost, D., Papademetris, X., Constable, R.T. (2012) Intrinsic brain connectivity related to age in young and middle aged adults. PloS one, 7:e44067.
4. Buckner, R.L., Sepulcre, J., Talukdar, T., Krienen, F.M., Liu, H., Hedden, T., Andrews-Hanna, J.R., Sperling, R.A., Johnson, K.A. (2009) Cortical Hubs Revealed by Intrinsic Functional Connectivity: Mapping, Assessment of Stability, and Relation to Alzheimer's Disease. The Journal of Neuroscience, 29:1860-1873.
5. Joshi, A., Scheinost, D., Okuda, H., Belhachemi, D., Murphy, I., Staib, L.H., Papademetris, X. (2011) Unified framework for development, deployment and robust testing of neuroimaging algorithms. Neuroinformatics, 9:69-84.
6. Li, C.S., Sinha, R. (2008) Inhibitory control and emotional stress regulation: neuroimaging evidence for frontal-limbic dysfunction in psycho-stimulant addiction. Neurosci Biobehav Rev, 32:581-97.
7. Potenza, M.N., Hong, K.I., Lacadie, C.M., Fulbright, R.K., Tuit, K.L., Sinha, R. (2012) Neural correlates of stress-induced and cue-induced drug craving: influences of sex and cocaine dependence. Am J Psychiatry, 169:406-14.
8. Rubinov, M., Sporns, O. (2010) Complex network measures of brain connectivity: Uses and interpretations. NeuroImage, 52:1059-1069.
9. Scheinost, D., Benjamin, J., Lacadie, C.M., Vohr, B., Schneider, K.C., Ment, L.R., Papademetris, X., Constable, R.T. (2012) The Intrinsic Connectivity Distribution: A Novel Contrast Measure Reflecting Voxel Level Functional Connectivity. Neuroimage.
10. Seo, D., Jia, Z., Lacadie, C.M., Tsou, K.A., Bergquist, K., Sinha, R. (2011) Sex differences in neural responses to stress and alcohol context cues. Hum Brain Mapp, 32:1998-2013.