Methods for Whole-Brain Comparisons of Resting State Functional Connectivity

Stephen J. Gotts Laboratory of Brain and Cognition NIMH/NIH Bethesda, MD

Functional Connectivity of Spontaneous Activity at Rest (i.e. "Resting State")

Functional Connectivity of Spontaneous Activity at Rest (i.e. "Resting State")

- very popular (easy and fast to administer)
- subjects passively view a fixation cross
- fluctuations in spontaneous activity (< .1 Hz) are correlated throughout the brain in a spatially restricted manner

Functional Connectivity of Spontaneous Activity at Rest (i.e. "Resting State")

- very popular (easy and fast to administer)
- subjects passively view a fixation cross
- fluctuations in spontaneous activity (< .1 Hz) are correlated throughout the brain in a spatially restricted manner

How can we do brain-wide testing in fMRI?

How can we do brain-wide testing in fMRI?

How can we do brain-wide testing in fMRI?

Typical methods for Task-based fMRI:

• Multiple regression to get Beta Weights per condition/subject

How can we do brain-wide testing in fMRI?

- Multiple regression to get Beta Weights per condition/subject
- Test Beta Weights in a voxel-wise manner across subjects (using t-tests, ANOVA, correlation, etc.)

How can we do brain-wide testing in fMRI?

- Multiple regression to get Beta Weights per condition/subject
- Test Beta Weights in a voxel-wise manner across subjects (using t-tests, ANOVA, correlation, etc.)
- threshold statistic and correct for voxel-wise comparisons using cluster-size from Monte Carlo simulations or FDR

How can we do brain-wide testing in fMRI?

- Multiple regression to get Beta Weights per condition/subject
- Test Beta Weights in a voxel-wise manner across subjects (using t-tests, ANOVA, correlation, etc.)
- threshold statistic and correct for voxel-wise comparisons using cluster-size from Monte Carlo simulations or FDR
- This holds for single whole-brain tests and can be adjusted for multiple tests on the same data using Bonferroni on the corrected p-value

How can we do brain-wide testing in fMRI?

Typical methods for Task-based fMRI:

- Multiple regression to get Beta Weights per condition/subject
- Test Beta Weights in a voxel-wise manner across subjects (using t-tests, ANOVA, correlation, etc.)
- threshold statistic and correct for voxel-wise comparisons using cluster-size from Monte Carlo simulations or FDR
- This holds for single whole-brain tests and can be adjusted for multiple tests on the same data using Bonferroni on the corrected p-value

But what to do for functional connectivity studies?

Brain-wide testing for functional connectivity?

Brain-wide testing for functional connectivity?

The number of comparisons explodes for voxel-wise tests. Every voxel with every voxel is a lot of tests for 40,000 voxels:

Brain-wide testing for functional connectivity?

The number of comparisons explodes for voxel-wise tests. Every voxel with every voxel is a lot of tests for 40,000 voxels:

40,000 x 39,999 / 2 \approx 8.0 x 10⁸ tests

Brain-wide testing for functional connectivity?

The number of comparisons explodes for voxel-wise tests. Every voxel with every voxel is a lot of tests for 40,000 voxels:

40,000 x 39,999 / 2 \approx 8.0 x 10⁸ tests $P < .05/(8.0 \times 10^8) = 6.125 \times 10^{-11}$ (Bonferroni)

Brain-wide testing for functional connectivity?

The number of comparisons explodes for voxel-wise tests. Every voxel with every voxel is a lot of tests for 40,000 voxels:

40,000 x 39,999 / 2 \approx 8.0 x 10⁸ tests $P < .05/(8.0 \times 10^8) = 6.125 \times 10^{-11}$ (Bonferroni)

False Discovery Rate might work (if most p-values are low), but not for more selective differences in typical sized datasets

Brain-wide testing for functional connectivity?

The number of comparisons explodes for voxel-wise tests. Every voxel with every voxel is a lot of tests for 40,000 voxels:

40,000 x 39,999 / 2 \approx 8.0 x 10⁸ tests $P < .05/(8.0 \times 10^8) = 6.125 \times 10^{-11}$ (Bonferroni)

False Discovery Rate might work (if most p-values are low), but not for more selective differences in typical sized datasets

Other options:

Brain-wide testing for functional connectivity?

The number of comparisons explodes for voxel-wise tests. Every voxel with every voxel is a lot of tests for 40,000 voxels:

40,000 x 39,999 / 2 \approx 8.0 x 10⁸ tests $P < .05/(8.0 \times 10^8) = 6.125 \times 10^{-11}$ (Bonferroni)

False Discovery Rate might work (if most p-values are low), but not for more selective differences in typical sized datasets

Other options:

- Predefined Regions of Interest
	- but might not capture the full picture

Brain-wide testing for functional connectivity?

The number of comparisons explodes for voxel-wise tests. Every voxel with every voxel is a lot of tests for 40,000 voxels:

40,000 x 39,999 / 2 \approx 8.0 x 10⁸ tests $P < .05/(8.0 \times 10^8) = 6.125 \times 10^{-11}$ (Bonferroni)

False Discovery Rate might work (if most p-values are low), but not for more selective differences in typical sized datasets

Other options:

- Predefined Regions of Interest
	- but might not capture the full picture
- Methods that decompose the data into smaller numbers of elements, such as **ICA**
	- requires some assumptions about the nature of the data

Two different whole-brain approaches that are more purely statistical (based in cluster-size correction), with fewer *a priori* assumptions about network structure:

Two different whole-brain approaches that are more purely statistical (based in cluster-size correction), with fewer *a priori* assumptions about network structure:

• Using average "connectedness" (centrality)

Two different whole-brain approaches that are more purely statistical (based in cluster-size correction), with fewer *a priori* assumptions about network structure:

- Using average "connectedness" (centrality)
- Testing every voxel as a seed (without averaging)

Compress the all-to-all voxels problem into a single map of "connectedness" for each subject (per condition)

Compress the all-to-all voxels problem into a single map of "connectedness" for each subject (per condition)

** a la Bob Cox and his AFNI group*

Compress the all-to-all voxels problem into a single map of "connectedness" for each subject (per condition)

Group Average Connectedness (per condition):

Compress the all-to-all voxels problem into a single map of "connectedness" for each subject (per condition)

Compress the all-to-all voxels problem into a single map of "connectedness" for each subject (per condition)

Pro: Preserves a lot of the spatial resolution in the data, Regardless of the group comparison, has a shot at finding "under" or "over-connected" voxels

Con: Might miss more spatially restricted effects and mixtures of under/over-connection

Example: Autism (ASD) vs. Typically Developing (TD)

Fractionation of social brain circuits in autism spectrum disorders

Stephen J. Gotts,¹ W. Kyle Simmons,² Lydia A. Milbury,¹ Gregory L. Wallace,¹ Robert W. Cox³ and Alex Martin¹

31 High-Functioning ASD adolescents

- Using DSM-IV criteria + ADI, ADOS
- "Triad" of impairments:
	- Impaired social functioning
	- Restricted interests/repetitive behaviors
	- Language/communication impairments

29 Typically Developing (TD) controls

31 High-Functioning ASD adolescents

- Using DSM-IV criteria + ADI, ADOS
- "Triad" of impairments:
	- Impaired social functioning
	- Restricted interests/repetitive behaviors
	- Language/communication impairments

29 Typically Developing (TD) controls

Groups matched on:

AGE:~17 (12-24) IQ: ~113 (85-143)

Sex 95% male subjects

31 High-Functioning ASD adolescents

- Using DSM-IV criteria + ADI, ADOS
- "Triad" of impairments:
	- Impaired social functioning
	- Restricted interests/repetitive behaviors
	- Language/communication impairments

29 Typically Developing (TD) controls

Groups matched on:

AGE:~17 (12-24) IQ: ~113 (85-143) Sex 95% male subjects

Scanned at rest with 3.5 sec TR for 8 min 10 sec with 1.7 x 1.7 x 3 voxels

How is functional connectivity altered in ASD ?

How is functional connectivity altered in ASD ?

• Generalized disruption of all 'circuits' ?

 ... or System-specific disruption ? (e.g. circuits involved in social processing)

How is functional connectivity altered in ASD ?

• Generalized disruption of all 'circuits' ?

 ... or System-specific disruption ? (e.g. circuits involved in social processing)

• Increase in Local Interactions ? (**)
The "Social Brain" (a la Brothers, 1990; Frith & Frith, 2007; Adolphs, 2009)

Using Group Connectedness to Find Seeds

ASD TD

Using Group Connectedness to Find Seeds

TD - ASD

Using Group Connectedness to Find Seeds

TD - ASD

Seeds: • p<.05 • at least 100 voxels

Yields 14 Seeds

Seeds + Seed Tests --> 27 Total Regions of Interest

Seeds + Seed Tests --> 27 Total Regions of Interest

Seeds + Seed Tests --> 27 Total Regions of Interest

How do these areas relate to each other ?

h

Cluster 1: Social inference/affective Cluster 2: Control/selection-retrieval Cluster 3: Social perception Form / motion

Back to ROI-ROI Correlation Matrices

Is this clinically relevant?

Cluster 2: Language / communication

Cluster 1: Social inference/affective

'Functional decoupling'

Cluster 3: Social perception Form / action

Correlations of **Social Responsiveness Scale (SRS)** ROI x ROI correlations in ASD sample alone (N=29)

Correlations of **Social Responsiveness Scale (SRS)** with Connectedness in ASD sample alone (N=29)

Summary for ASD Study

- At least for high-functioning ASD subjects, the largest differences in correlation were concentrated among regions of the 'social brain'
- We observed a fractionation of social brain circuits into two parts
- Social/affective component (Cluster 1) was 'functionally' decoupled from language and visuomotor components

Applying the same method to Childhood Onset Schizophrenia (vs. Typ. Developing)

Collaboration with: Becky Berman Harrison McAdams Nitin Gogtay Judy Rapoport *et al.*

Red cluster:

Social-cognitive

Green cluster:

Sensorimotor

 \circ spatial overlap with seed-based symptom correlation analyses (p<.05, corrected by cluster size)

 \circ spatial overlap with seed-based symptom correlation analyses Group t-tests: $(p < .05,$ corrected by cluster size)

Increased Resting Correlations in Primary Lateral Sclerosis (PLS)

Collaboration with Mary Kay Floeter (NINDS) and Avner Meoded (Johns Hopkins):

Cerebro-cerebellar connectivity is increased in primary lateral sclerosis

Avner Meoded^{a, 1}, Arthur E. Morrissette^{a, 2}, Rohan Katipally^{a, 3}, Olivia Schanz^a, Stephen J. Gotts^b, Mary Kay Floeter^{a,*}

a National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

^b National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA

Increased Resting Correlations in Primary Lateral Sclerosis (PLS)

Collaboration with Mary Kay Floeter (NINDS) and Avner Meoded (Johns Hopkins):

Controls PLS PLS-Control

Group tests of whole-brain connectedness, along with subsequent seed tests, can detect brain regions that are related to behavioral functions of interest (e.g. social ability in Autism)

Group tests of whole-brain connectedness, along with subsequent seed tests, can detect brain regions that are related to behavioral functions of interest (e.g. social ability in Autism)

However, it's still not clear that everything is being detected:

Group tests of whole-brain connectedness, along with subsequent seed tests, can detect brain regions that are related to behavioral functions of interest (e.g. social ability in Autism)

However, it's still not clear that everything is being detected:

• problem of mixtures that cancel

Group tests of whole-brain connectedness, along with subsequent seed tests, can detect brain regions that are related to behavioral functions of interest (e.g. social ability in Autism)

However, it's still not clear that everything is being detected:

- problem of mixtures that cancel
- spatially restricted effects can fail to be detected

Testing Every Voxel as a Seed

Testing Every Voxel as a Seed

Eliminates the averaging approach, but can take a long time (\sim 2 weeks on a fast desktop with 32 GB of RAM)

Testing Every Voxel as a Seed

Eliminates the averaging approach, but can take a long time (\degree 2 weeks on a fast desktop with 32 GB of RAM)

Basic Approach: Adjust Monte Carlo simulations of cluster-size to handle testing of all seeds
Eliminates the averaging approach, but can take a long time (\degree 2 weeks on a fast desktop with 32 GB of RAM)

Basic Approach: Adjust Monte Carlo simulations of cluster-size to handle testing of all seeds

• Generate random data for N simulated subjects

Eliminates the averaging approach, but can take a long time (\degree 2 weeks on a fast desktop with 32 GB of RAM)

- Generate random data for N simulated subjects
- Spatially blur to match average smoothness of actual data

Eliminates the averaging approach, but can take a long time (\degree 2 weeks on a fast desktop with 32 GB of RAM)

- Generate random data for N simulated subjects
- Spatially blur to match average smoothness of actual data
- Conduct correlation tests using every voxel as a seed, keeping track of the largest cluster ever detected surviving a particular voxel-wise p-value

Eliminates the averaging approach, but can take a long time (\degree 2 weeks on a fast desktop with 32 GB of RAM)

- Generate random data for N simulated subjects
- Spatially blur to match average smoothness of actual data
- Conduct correlation tests using every voxel as a seed, keeping track of the largest cluster ever detected surviving a particular voxel-wise p-value
- Repeat many times (e.g. 5000 iterations) to determine the cluster size needed for P<.05 FWE

Eliminates the averaging approach, but can take a long time (\degree 2 weeks on a fast desktop with 32 GB of RAM)

- Generate random data for N simulated subjects
- Spatially blur to match average smoothness of actual data
- Conduct correlation tests using every voxel as a seed, keeping track of the largest cluster ever detected surviving a particular voxel-wise p-value
- Repeat many times (e.g. 5000 iterations) to determine the cluster size needed for P<.05 FWE
- Do same tests on actual data using these critical thresholds to find corrected results

56 ASD, 62 TD, separated into two independent sets (Sets 1 and 2: 28 ASD, 31 TD each) that are matched for Motion and Age (P>.1 for all)

All voxel-wise t-tests also include Motion and Age as covariates (AFNI's 3dttest++, with common median centering)

Cluster-size Thresholds from Monte Carlo Simulations (5000 iterations)

Analysis Mask (85% of Both ASD/TD Groups)

Set 1 (28 ASD, 31 TD)

Seed Voxels involved in significant differences for which **TD > ASD** (ranging from P<.05 down to P<.00005, corrected):

Set 1 (28 ASD, 31 TD) Seed Voxels involved in significant differences for which **TD > ASD** (ranging from P<.05 down to P<.00005, corrected):

Set 1 (28 ASD, 31 TD)

Seed Voxels involved in significant differences for which **ASD > TD** (ranging from P<.05 down to P<.00005, corrected):

Set 1 (28 ASD, 31 TD) Seed Voxels involved in significant differences for which **ASD > TD** (ranging from P<.05 down to P<.00005, corrected):

What is the relationship to Connectedness comparisons?

... and the previously reported results? (Brain 2012)

(31 ASD, 29 TD)

27 ROIs (ANATICOR)

Larger NIMH Dataset

(56 ASD, 62 TD)

Connectedness Tests (TD-ASD), P<.05, uncorrected (>100 voxels)

(31 ASD, 29 TD)

27 ROIs (ANATICOR)

Larger NIMH Dataset

(56 ASD, 62 TD)

Connectedness Tests (TD-ASD), P<.05, uncorrected (>100 voxels)

(31 ASD, 29 TD)

27 ROIs (ANATICOR)

Larger NIMH Dataset

(56 ASD, 62 TD)

Connectedness Tests (TD-ASD), P<.05, corrected

(31 ASD, 29 TD)

Larger NIMH Dataset

(56 ASD, 62 TD)

Connectedness Tests (TD-ASD), Replication Across Two Sets, P<.05, corrected

(31 ASD, 29 TD)

Larger NIMH Dataset

(56 ASD, 62 TD)

Testing All Voxels as Seeds (TD-ASD), Replication Across Two Sets, P<.05, corrected

(31 ASD, 29 TD)

Larger NIMH Dataset

(56 ASD, 62 TD)

Testing All Voxels as Seeds (TD-ASD), Replication Across Two Sets, ASD>TD

Brain 2012 ROIs (31 ASD, 29 TD)

Larger NIMH Dataset

(56 ASD, 62 TD)

Testing All Voxels as Seeds (TD-ASD), Replication Across Two Sets, P<.05, corrected

(31 ASD, 29 TD)

27 ROIs (ANATICOR)

Larger NIMH Dataset

(56 ASD, 62 TD)

Connectedness Tests (TD-ASD), P<.05, corrected

• Massive voxel-wise testing of functional connectivity data is indeed possible, with robust replication across independent datasets

- Massive voxel-wise testing of functional connectivity data is indeed possible, with robust replication across independent datasets
- Voxel-wise seed testing does a better job than connectedness tests at identifying locations with mixed results (TD>ASD and ASD>TD), although results are not radically different

- Massive voxel-wise testing of functional connectivity data is indeed possible, with robust replication across independent datasets
- Voxel-wise seed testing does a better job than connectedness tests at identifying locations with mixed results (TD>ASD and ASD>TD), although results are not radically different
- Such tests do not require a priori assumptions about common network structure in control and clinical groups (as group ICA methods commonly do)

- Massive voxel-wise testing of functional connectivity data is indeed possible, with robust replication across independent datasets
- Voxel-wise seed testing does a better job than connectedness tests at identifying locations with mixed results (TD>ASD and ASD>TD), although results are not radically different
- Such tests do not require a priori assumptions about common network structure in control and clinical groups (as group ICA methods commonly do)
- Searches are possible for any type of test statistic for which p-values can be calculated (e.g. correlation with behavioral measures, more complex ANOVAs, etc.)

Acknowledgements:

Section on Cognitive Neuropsychology, LBC (NIMH)

Alex Martin, Chief Kyle Simmons (now at *LIBR*, Tulsa) Lydia Milbury Greg Wallace

Scientific and Statistical Computing Core (NIMH)

Bob Cox, Chief Ziad Saad Gang Chen