DPABI: Quality Control, Statistical Analysis and Results Viewing

Chao-Gan YAN, Ph.D.
严超赣
ycg.yan@gmail.com
http://rfmri.org
Institute of Psychology, Chinese Academy of Sciences

1

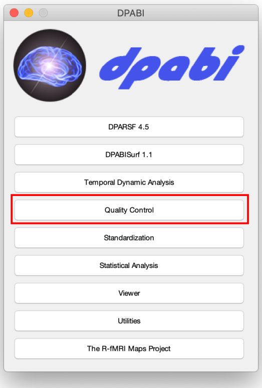
1

Outline

- Quality Control
- Statistical Analysis
- Results Viewing

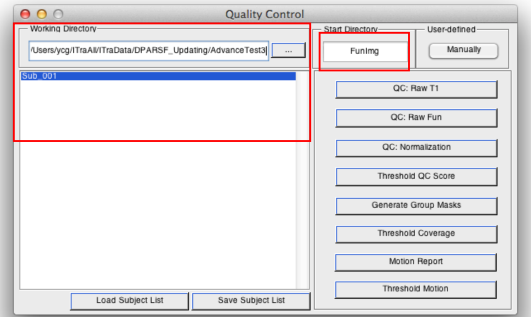
2

2



3

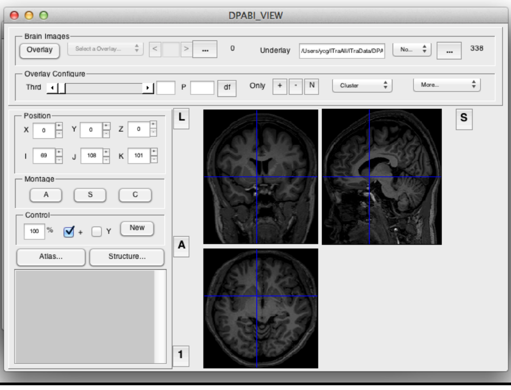
Quality Control



4

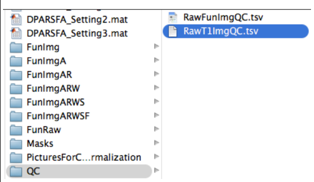
4

Quality Control



5

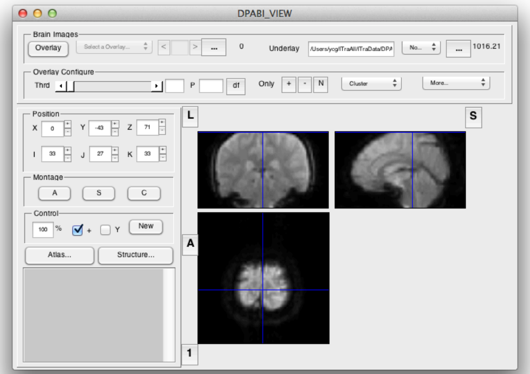
Quality Control



	A	B	C	D
1	Subject ID	QC Score	QC Comment	
2	Sub_001	5	Very Good!	
3				

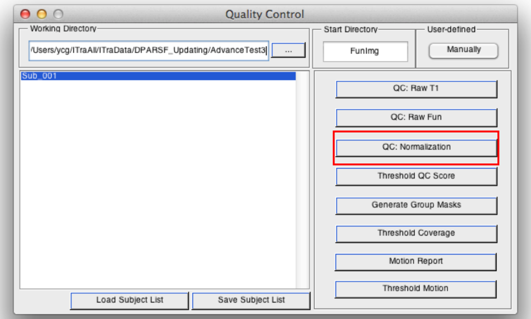
6

Quality Control



7

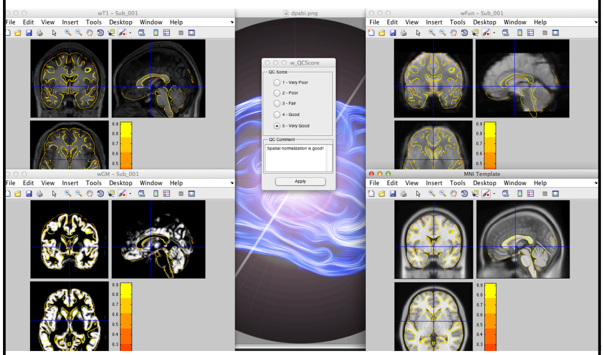
Quality Control



8

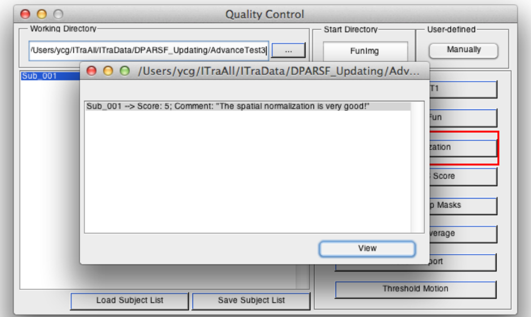
8

Quality Control



9

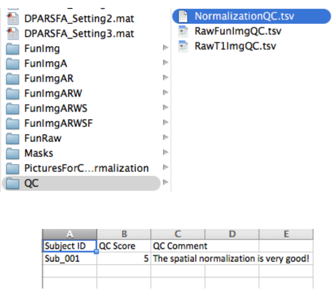
Quality Control



10

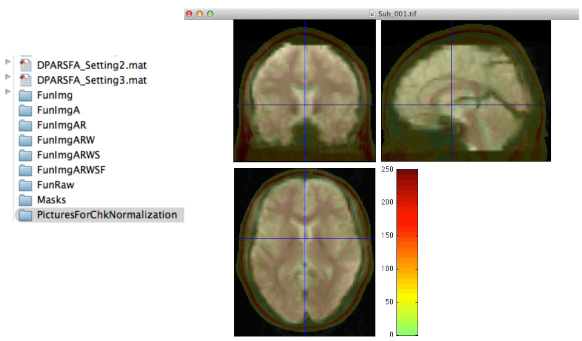
10

Quality Control

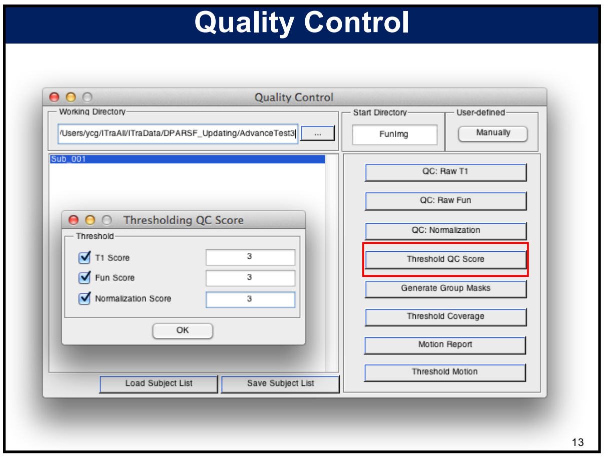


11

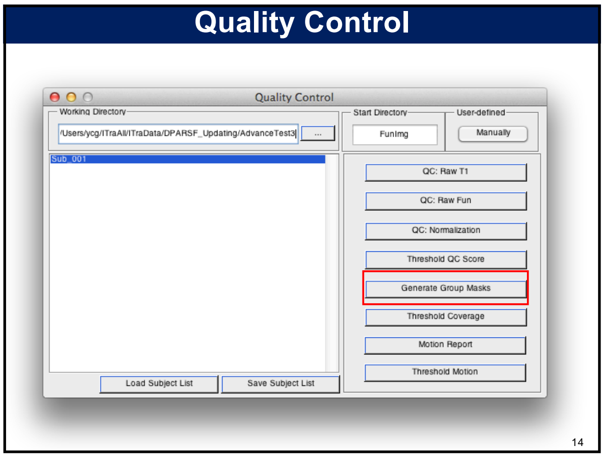
Quality Control



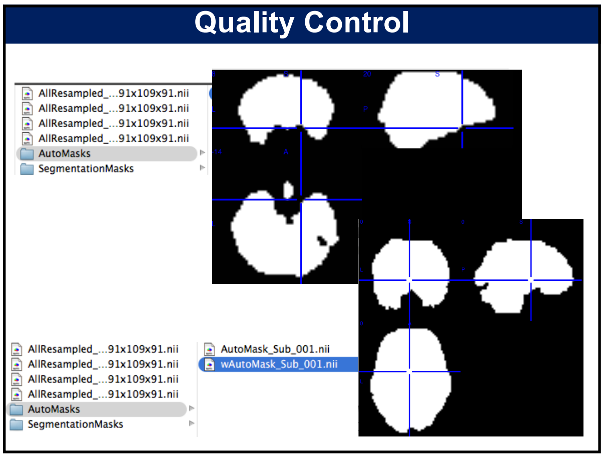
12



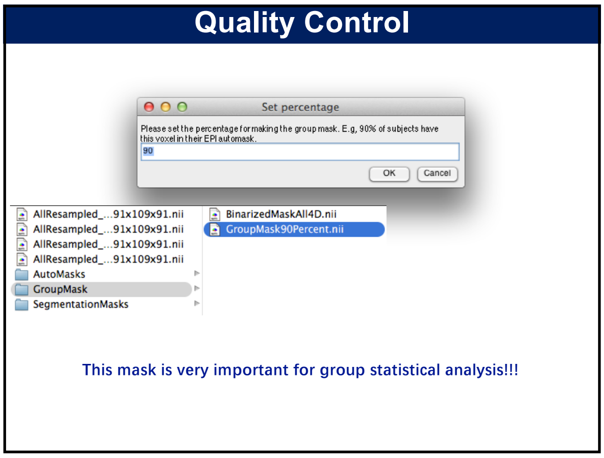
13



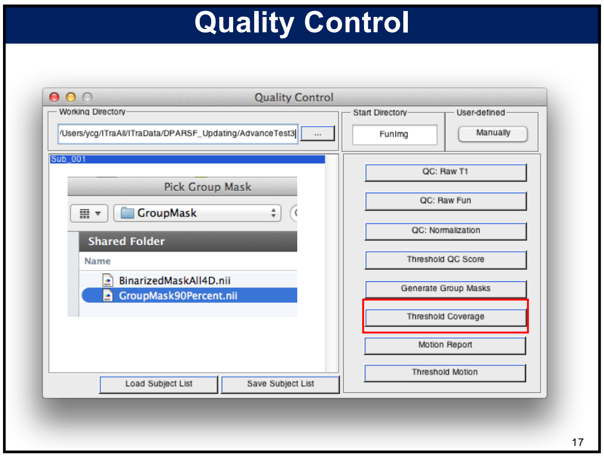
14



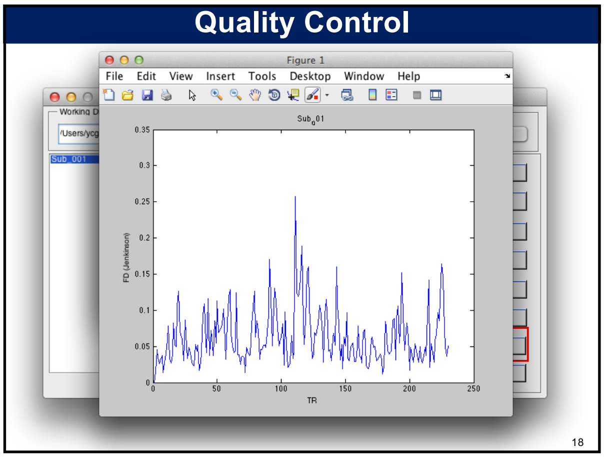
15



16

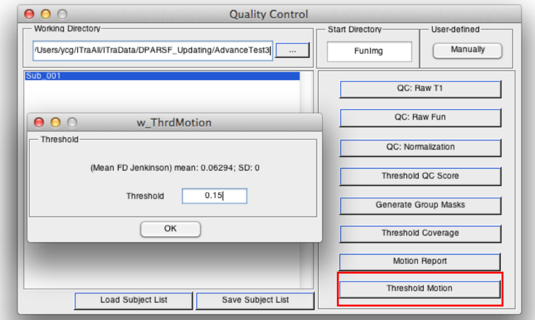


17



18

Quality Control

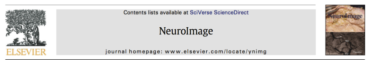


19

Quality Control

- Using the visual inspection step within DPARSF, subjects showing severe head motion in the T1 image and subjects showing extremely poor coverage in the functional images, as well as subjects showing bad registration were excluded
- Subjects with overlap with the group mask (voxels present at least 90% of the participants) less than 2*SD under the group mean overlap (threshold: 92.2%) were excluded
- Subjects with motion (Mean FD Jenkinson greater than 2*SD above the group mean motion (threshold: 0.192) were excluded

Yan et al., 2013, Neuroimage


Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Full Length Article

Standardizing the intrinsic brain: Towards robust measurement of inter-individual variation in 1000 functional connectomes

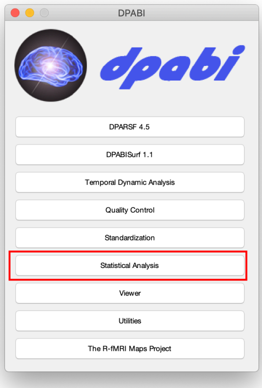
Chao-Gan Yan [a,b,c], R. Cameron Craddock [a,b], Xi-Nian Zuo [d], Yu-Feng Zang [e], Michael P. Milham [a,b,*]

20

Outline

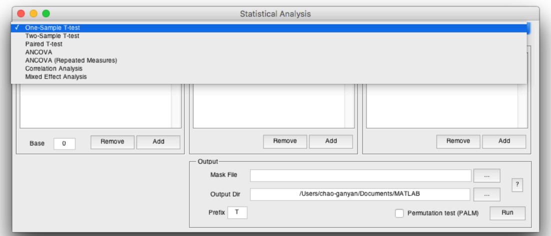
- Quality Control
- ➡ Statistical Analysis
- Results Viewing

21



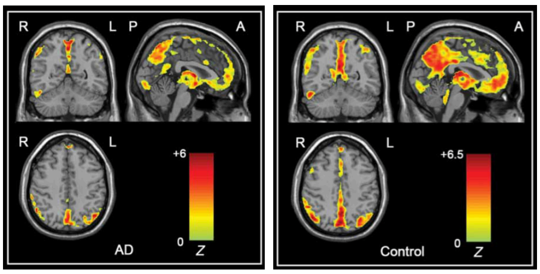
22

Statistical Analysis



23

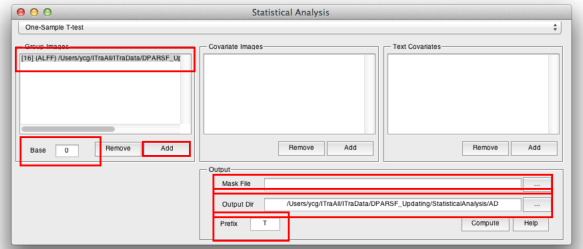
One-Sample T-Test



Wang[#], Yan[#] et al., 2011, Hum Brain Mapp

24

One-Sample T-Test



0 for z* images
1 for m* images

T Statistic Image

Group Mask

25

Two-Sample T-Test

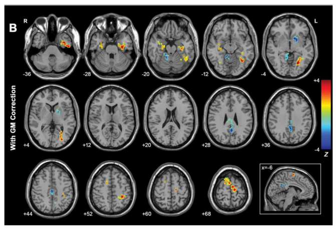


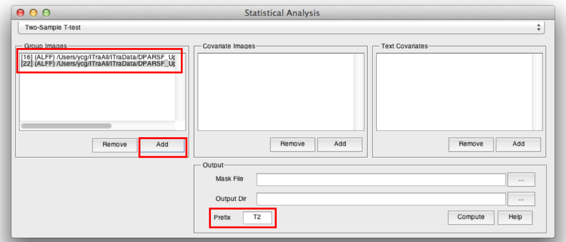
Figure 3.

A: *Z*-statistical difference maps between the AD patients and healthy eldely (without GM correction). The AD patients showed significantly decreased ALFF in the bilateral PCC/PCu, left LN, left OFG, and right ALC, and increased ALFF in the bilateral PHG, bilateral Hip, bilateral SFG, bilateral SMA, left FG, left PoCG, left ITG, and left STG. For the details of the regions, see Table III. **B**: *Z*-statistical difference maps between the AD patients and healthy elderly (with GM correction). The AD patients showed significantly decreased ALFF in the bilateral PCC/PCu, left LN, and right ALC, and increased ALFF in the bilateral PHG, bilateral Hip, bilateral SFG, bilateral SMA, left FG, left PoCG, left ITG, and left STG. For the details of the regions, see Table III. The statistical threshold was set at $|Z| > 1.96$ ($P < 0.05$) and cluster size $>1,404$ mm^3, which corresponded to a corrected $P < 0.05$. Of note, we showed the two-sample *t*-tests results within a mask showing significant group differences in the ANOVA analysis (Fig. 2A). R, right; L, left. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Wang#, Yan# et al., 2011, Hum Brain Mapp

26

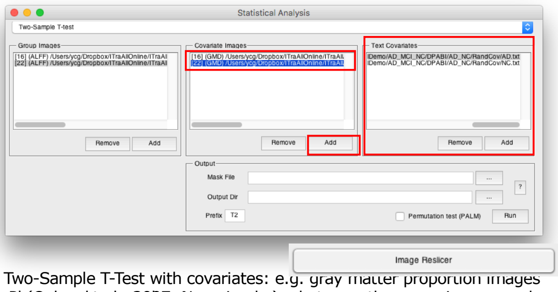
Two-Sample T-Test



T Statistic Image: positive corresponds to the mean of Group 1 is greater than the mean of Group 2

27

Two-Sample T-Test



Two-Sample T-Test with covariates: e.g. gray matter proportion images (Oakes et al., 2007, Neuroimage) Other covariate can be also specified as text files. (E.g. head motion (mean FD), age, sex etc.)

Please make sure the correspondence between the group images and the covariate images: order and voxel size

28

Paired T-Test

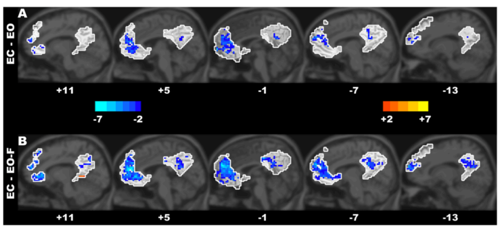
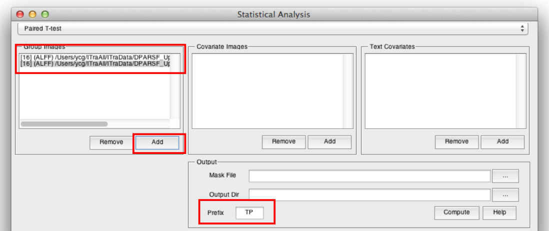


Figure 3. The between-condition differences of the ALFF within the DMN. The ALFF differences were found between the EC and EO conditions (A), and between the EC and EO-F conditions (B). The areas in the white contours denote the ROIs within the DMN. The numbers below the images refer to the x coordinates in the Talairach and Tournoux space. The statistical threshold was set at $|t|>2.093$ ($P<0.05$) and cluster size >486 mm^3, which corresponds to a corrected $P<0.05$.
doi:10.1371/journal.pone.0005743.g003

Yan et al., 2009. PLoS ONE

29

Paired T-Test



T Statistic Image

Condition 1 – Condition 2
Please make sure the correspondence

30

ANOVA or ANCOVA

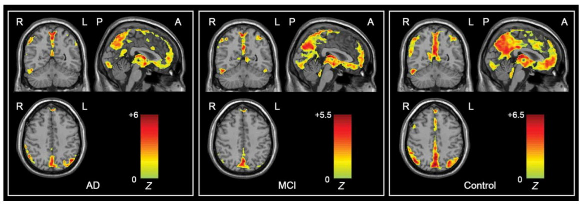


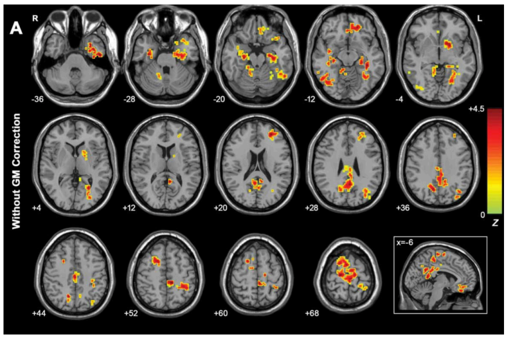
Figure 1.
Within-group ALFF maps within the AD, MCI, and healthy elderly control groups. Visual inspection indicated that the PCC and adjacent PCu had the highest ALFF values within each group and had different strengths among the three groups. The statistical threshold was set at $Z > 3.09$ ($P < 0.001$) and cluster size >189 mm^3, which corresponded to a corrected $P < 0.001$. R, right; L, left; P, posterior; A, anterior. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Wang#, Yan# et al., 2011, Hum Brain Mapp

31

31

ANOVA or ANCOVA

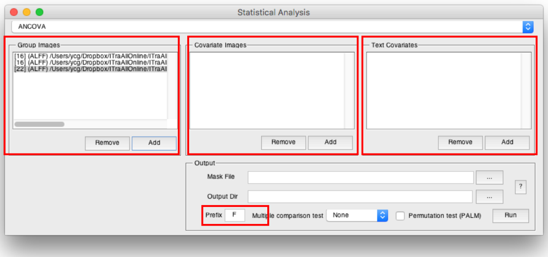


Wang#, Yan# et al., 2011, Hum Brain Mapp

32

32

ANOVA or ANCOVA



ANCOVA: e.g. gray matter proportion images (Oakes et al., 2007, Neuroimage)

Please make sure the correspondence between the group images and the covariate images: order and voxel size

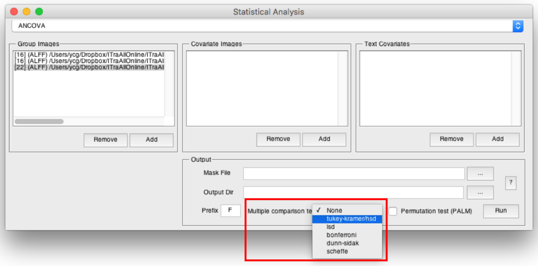
F Statistic Image

Other covariate can be also specified as text files. (E.g. age, brain size, IQ etc.)

33

33

ANOVA or ANCOVA



Post-hoc procedures: the corrected p values under a given control procedure for comparing group means of any pairs were calculated (e.g., through Studentized Range statistic for Tukey-Kramer correction) with the same route as MATLAB command multcompare. The p maps were then converted to Z maps according to the Normal inverse cumulative distribution function (norminv), with the sign of group mean differences applied.

34

34

ANOVA or ANCOVA

ANOVA F image

The difference of mean between groups

The corrected p of difference between groups

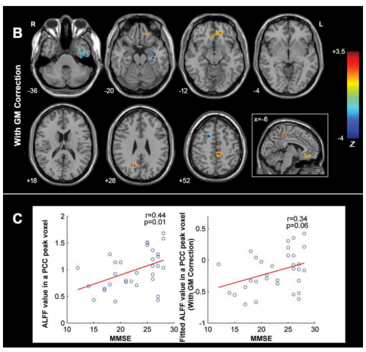
The corrected Z values of difference between groups, can be forwarded to further multiple comparison correction

Yan et al., 2016. Neuroinformatics

35

35

Correlation Analysis

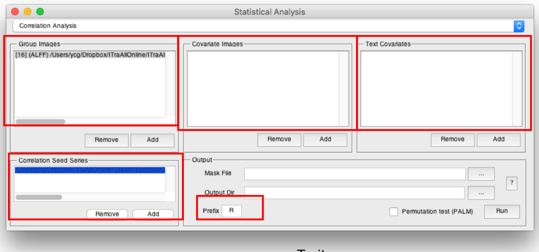


Wang#, Yan# et al., 2011, Hum Brain Mapp

36

36

Correlation Analysis

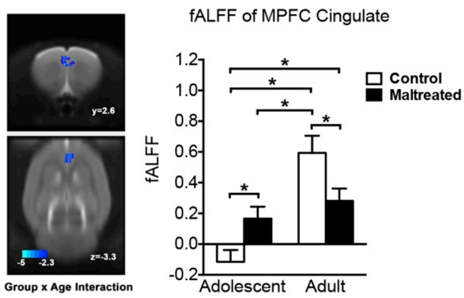


The imaging measure:
ALFF maps

Traits: e.g.
MMSE.txt
19
15
…

37

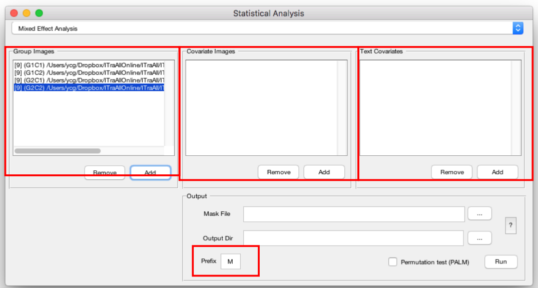
Mixed Effect Analysis



Yan et al., 2016. Translational Psychiatry

38

Mixed Effect Analysis



The imaging measure should be:
Group1Condition1
Group1Condition2
Group2Condition1
Group2Condition2

39

Mixed Effect Analysis

- *_ConditionEffect_T.nii - the T values of condition differences (corresponding to the first condition minus the second condition) (WithinSubjectFactor)
- *_Interaction_F.nii - the F values of interaction (BetweenSubjectFactor by WithinSubjectFactor)
- *_Group_TwoT.nii - the T values of group differences (corresponding to the first group minus the second group). Of note: the two conditions will be averaged first for each subject. (BetweenSubjectFactor)

40

Statistical Analysis

```
function [b_OLS_brain, t_OLS_brain, TF_ForContrast_brain, r_OLS_brain, Header] = y_GroupAnalysis_Image(DependentVolume,Predictor,OutputName,
% function [b_OLS_brain, t_OLS_brain, TF_ForContrast_brain, r_OLS_brain, Header] = y_GroupAnalysis_Image(DependentVolume,Predictor,OutputNam
% Perform regression analysis
% Input:
%   DependentVolume     -   4D data matrix (DimX*DimY*DimZ*DimTimePoints) or the directory of 3D image data file or the filename of one 4D d
%   Predictor - the Predictors M (subjects) by N (traits). SHOULD INCLUDE the CONSTANT column if needed. The program will not add constant c
%   OutputName - the output name. (should not have extention such as .img,.nii)
%   MaskFile - the mask file.
%   CovVolume  [optional] - 4D data matrix (DimX*DimY*DimZ*DimTimePoints) or the directory of image covariates, in which the files should be
%   Contrast [optional] - Contrast for T-test for F-test. 1*ncolX matrix.
%   TF_Flag [optional] - 'T' or 'F'. Specify if T-test or F-test need to be performed for the contrast
%   IsOutputResidual [optional] - 1: output the 4D residuals.
%                               - 0: don't output the 4D residuals
%   Header [optional] - If DependentVolume is given as a 4D Brain matrix, then Header should be designated.

% Output:
%   OutputName_b.nii, OutputName_T.nii     - beta and t value files results
%   OutputName_Residual.nii (optional)     - Residual files
%___________________________________________________________________________
% Written by YAN Chao-Gan 120823.
% The Nathan Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA
% Child Mind Institute, 445 Park Avenue, New York, NY 10022, USA
% The Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, New York University Child Study Center, New York, NY 10016, USA
% ycg.yan@gmail.com
```

{DPABI_Dir}/StatisticalAnalysis/y_GroupAnalysis_Image.m

41

Statistical Analysis

42

Statistical Analysis

```
function [b_OLS_brain, t_OLS_brain, TF_ForContrast_brain, r_OLS_brain, Header] = y_GroupAnalysis_Image(DependentVolume,Predictor,OutputName,
% function [b_OLS_brain, t_OLS_brain, TF_ForContrast_brain, r_OLS_brain, Header] = y_GroupAnalysis_Image(DependentVolume,Predictor,OutputNam
% Perform regression analysis
% Input:
%   DependentVolume     -   4D data matrix (DimX*DimY*DimZ*DimTimePoints) or the directory of 3D image data file or the filename of one 4D d
%   Predictor - the Predictors M (subjects) by N (traits). SHOULD INCLUDE the CONSTANT column if needed. The program will not add constant c
%   OutputName - the output name. (should not have extention such as .img,.nii)
%   MaskFile - the mask file.
%   CovVolume  [optional] - 4D data matrix (DimX*DimY*DimZ*DimTimePoints) or the directory of image covariates, in which the files should be
%   Contrast [optional] - Contrast for T-test for F-test. 1*ncolX matrix.
%   TF_Flag [optional] - 'T' or 'F'. Specify if T-test or F-test need to be performed for the contrast
%   IsOutputResidual [optional] - 1: output the 4D residuals.
%                    - 0: don't output the 4D residuals
%   Header [optional] - If DependentVolume is given as a 4D Brain matrix, then Header should be designated.

% Output:
%   OutputName_b.nii, OutputName_T.nii     - beta and t value files results
%   OutputName_Residual.nii (optional)     - Residual files
%___________________________________________________________________________
% Written by YAN Chao-Gan 120823.
% The Nathan Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA
% Child Mind Institute, 445 Park Avenue, New York, NY 10022, USA
% The Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, New York University Child Study Center, New York, NY 10016, USA
% ycg.yan@gmail.com
```

{DPABI_Dir}/StatisticalAnalysis/y_GroupAnalysis_Image.m

Smoothness estimation based on the 4D residual is built in this function!!!

43

Statistical Analysis

http://rfmri.org/DemoData {Download}/ProcessingDemoData/StatisticalDemo/AD_MCI_NC/

ALFF: AD – NC Two Sample T Test:

- Applied smooth kernel in preprocessing: [4 4 4]
- Smooth kernel estimated on 4D residual: [6.77 6.88 6.71]
- Smooth kernel estimated on statistical image (T to Z, as in easythresh): [6.90 7.33 6.94]

ReHo: AD – NC Two Sample T Test:

- Applied smooth kernel in preprocessing: [4 4 4]
- Smooth kernel estimated on 4D residual: [8.10 8.50 7.93]
- Smooth kernel estimated on statistical image (T to Z, as in easythresh): [8.33 8.94 8.24]

Thus, only using smooth kernel applied in preprocessing is NOT sufficient!!!

44

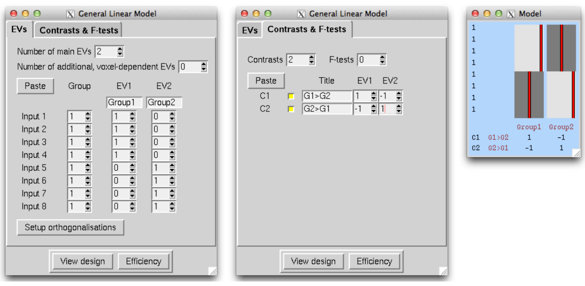
Statistical Analysis


```
function [b_OLS_brain, t_OLS_brain, TF_ForContrast_brain, r_OLS_brain, Header] = y_GroupAnalysis_Image(DependentVolume,Predictor,OutputName,
% function [b_OLS_brain, t_OLS_brain, TF_ForContrast_brain, r_OLS_brain, Header] = y_GroupAnalysis_Image(DependentVolume,Predictor,OutputNam
% Perform regression analysis
% Input:
%   DependentVolume     -   4D data matrix (DimX*DimY*DimZ*DimTimePoints) or the directory of 3D image data file or the filename of one 4D d
%   Predictor - the Predictors M (subjects) by N (traits). SHOULD INCLUDE the CONSTANT column if needed. The program will not add constant c
%   OutputName - the output name. (should not have extention such as .img,.nii)
%   MaskFile - the mask file.
%   CovVolume  [optional] - 4D data matrix (DimX*DimY*DimZ*DimTimePoints) or the directory of image covariates, in which the files should be
%   Contrast [optional] - Contrast for T-test for F-test. 1*ncolX matrix.
%   TF_Flag [optional] - 'T' or 'F'. Specify if T-test or F-test need to be performed for the contrast
%   IsOutputResidual [optional] - 1: output the 4D residuals.
%                    - 0: don't output the 4D residuals
%   Header [optional] - If DependentVolume is given as a 4D Brain matrix, then Header should be designated.

% Output:
%   OutputName_b.nii, OutputName_T.nii     - beta and t value files results
%   OutputName_Residual.nii (optional)     - Residual files
%___________________________________________________________________________
% Written by YAN Chao-Gan 120823.
% The Nathan Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA
% Child Mind Institute, 445 Park Avenue, New York, NY 10022, USA
% The Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, New York University Child Study Center, New York, NY 10016, USA
% ycg.yan@gmail.com
```


45

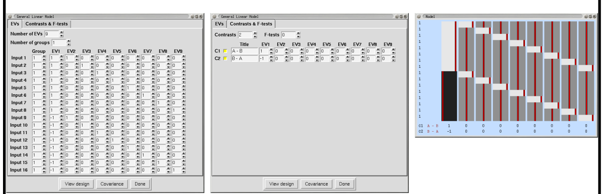
Statistical Analysis



http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM

46

Statistical Analysis



http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/GLM

47

Statistical Analysis

Multiple Comparison Correction

48

Multiple Comparison Correction

Eklund

49

49

Multiple Comparison Correction

… I estimate about 15,000 papers use cluster size inference with correction for multiple testing; of these, around 3,500 use a CDT of P=0.01…So, are we saying 3,500 papers are "wrong"? It depends….

-- Thomas Nichols
July 06, 2016

Correction

NEUROSCIENCE, STATISTICS
Correction for "Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates," by Anders Eklund, Thomas E. Nichols, and Hans Knutsson, which appeared in issue 28, July 12, 2016, of *Proc Natl Acad Sci USA* (113:7900–7905; first published June 28, 2016; 10.1073/pnas.1602413113).

The authors note that on page 7900, in the Significance Statement, lines 9–11, "These results question the validity of some 40,000 fMRI studies and may have a large impact on the interpretation of neuroimaging results" should instead appear as "These results question the validity of a number of fMRI studies and may have a large impact on the interpretation of weakly significant neuroimaging results."

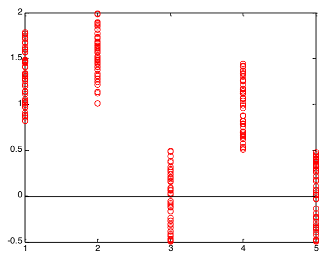
Additionally, the authors note that on page 7904, left column, fifth full paragraph, lines 1–3, "It is not feasible to redo 40,000 fMRI studies, and lamentable archiving and data-sharing practices mean most could not be reanalyzed either" should instead appear as "Due to lamentable archiving and data-sharing practices, it is unlikely that problematic analyses can be redone."

These errors do not affect the conclusions of the article. The online version has been corrected.

www.pnas.org/cgi/doi/10.1073/pnas.1612033113

50

Multiple Comparison Correction



P=0.05 P=0.05 P=0.05 P=0.05 P=0.05

0.95 0.95 0.95 0.95 0.95

0.95^5=0.774

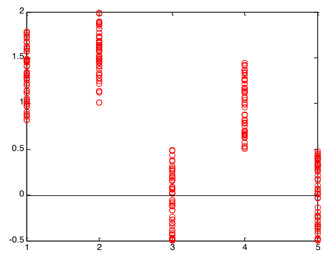
Probability of not getting a false positive result: 1 - 0.05 = 0.95 (for each of the five tests)

51

51

Multiple Comparison Correction

Bonferroni correction: p=0.05/5=0.01



0.01 0.01 0.01 0.01 0.01

52

Multiple Comparison Correction

- •False Discovery Rates (FDR) correction
- Family-Wise Error (FWE) correction
  - Bonferroni correction: 0.05/5=0.01
  - Gaussian Random Field theory correction
  - Monte Carlo simulations (AlphaSim)
  - Threshold-Free Cluster Enhancement
  - Permutation test

53

FDR Theory

Number of errors committed when testing m null hypotheses

	Declared non-significant	Declared significant	Total
True null hypotheses	U	V	m_0
Non-true null hypotheses	T	S	$m - m_0$
	$m - $ R	R	m

• False discovery rate Q_e=E(V/(V+S))=E(V/R)

Benjamini and Hochberg, 1995, Journal of the Royal Statistical Society

54

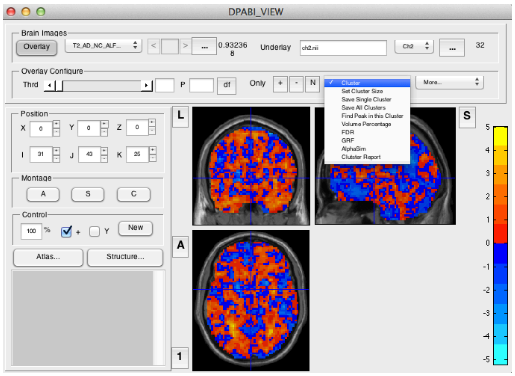
FDR Theory

• Let $H_1, \cdots, H_m$ be the null hypotheses and $P_1, \cdots, P_m$ their corresponding p-values. Order these values in increasing order and denote them by $P_{(1)}, \cdots, P_{(m)}$. For a given q, find the largest k such that $P_{(k)} \leqq kq/m$.

•Then reject (i.e. declare positive) all $H_{(i)}$ for i = 1, ⋯, k.

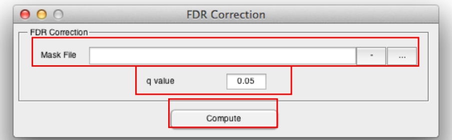
55

FDR Theory



56

FDR Theory

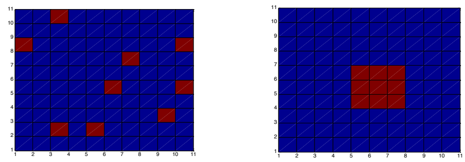


57

Multiple Comparison Correction

Gaussian Random Field Theory Correction

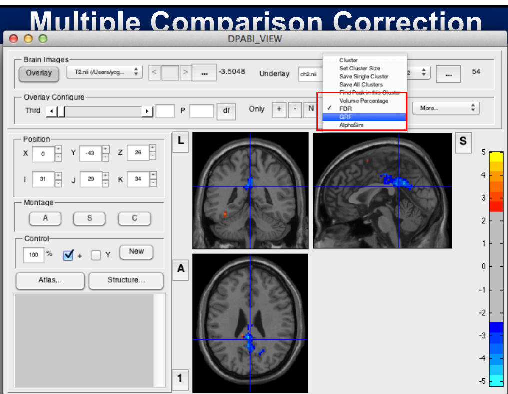
Monte Carlo simulations (AlphaSim)



?

58

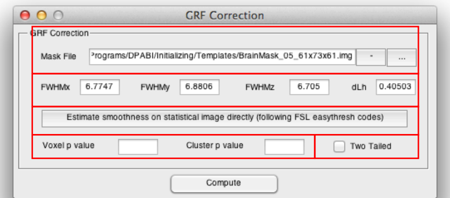
Multiple Comparison Correction



59

59

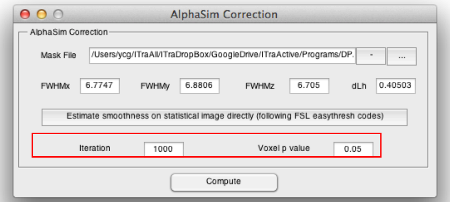
Multiple Comparison Correction



Voxel Z > 2.3, Cluster P < 0.05, Two One-Tailed Corrections: equivalent to
Voxel P < 0.0214, Cluster P < 0.1, Two Tailed.

60

Multiple Comparison Correction



61

Multiple Comparison Correction

Cl Size	Frequency	Cum Prop	p/Voxel	Max Freq	Alpha
1	235971	0.619898	0.009613	0	1.000000
2	76150	0.819945	0.006282	0	1.000000
3	32297	0.904789	0.004131	0	1.000000
4	15940	0.946664	0.002763	0	1.000000
5	8476	0.968930	0.001863	0	1.000000
6	4786	0.981503	0.001265	1	1.000000
7	2767	0.988772	0.000860	19	0.999000
8	1606	0.992991	0.000586	51	0.980000
9	1011	0.995647	0.000405	127	0.929000
10	585	0.997184	0.000276	132	0.802000
11	391	0.998211	0.000194	172	0.670000
12	236	0.998831	0.000133	146	0.498000
13	164	0.999262	0.000093	107	0.352000
14	98	0.999519	0.000063	78	0.245000
15	69	0.999701	0.000043	61	0.167000
16	37	0.999798	0.000029	30	0.106000
17	22	0.999856	0.000020	22	0.076000
18	22	0.999913	0.000015	21	0.054000
19	11	0.999942	0.000010	11	0.033000
20	7	0.999961	0.000007	7	0.022000
21	5	0.999974	0.000005	5	0.015000
22	5	0.999987	0.000003	5	0.010000
23	4	0.999997	0.000002	4	0.005000
24	1	1.000000	0.000000	1	0.001000

62

Threshold-Free Cluster Enhancement (TFCE)

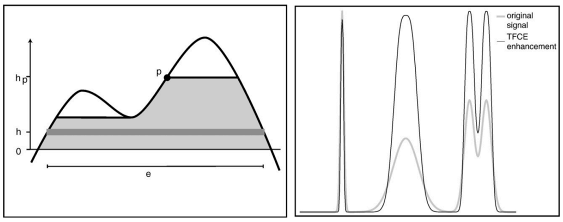
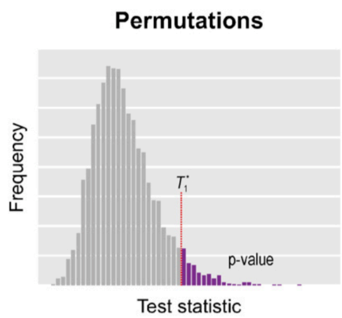


Fig. 1. Illustration of the TFCE approach. Left: the TFCE score at voxel *p* is given by the sum of the scores of all incremental supporting sections (one such is shown as the dark-grey band) within the area of "support" of *p* (light grey). The score for each section is a simple function of its height *h* and extent *e*. Right: example input image and TFCE-enhanced output. The input contains a focal, high signal, a much more spatially extended, lower, signal and a pair of overlapping signals of intermediate extent and height. The TFCE output has the same maximal values for all three cases, and preserves the distinct local maxima in the third case.

Smith et al., 2009. Neuroimage

63

Permutation Test



Winkler et al., 2016. Neuroimage

64

Multiple Comparison Correction

PNAS

CrossMark
click for updates

Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates

Anders Eklund[a,b,c,1], Thomas E. Nichols[d,e], and Hans Knutsson[a,c]

[a]Division of Medical Informatics, Department of Biomedical Engineering, Linköping University, S-581 85 Linköping, Sweden; [b]Division of Statistics and Machine Learning, Department of Computer and Information Science, Linköping University, S-581 83 Linköping, Sweden; [c]Center for Medical Image Science and Visualization, Linköping University, S-581 83 Linköping, Sweden; [d]Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom; and [e]WMG, University of Warwick, Coventry CV4 7AL, United Kingdom

Edited by Emery N. Brown, Massachusetts General Hospital, Boston, MA, and approved May 17, 2016 (received for review February 12, 2016)

Eklund et al., 2016. PNAS

65

Multiple Comparison Correction

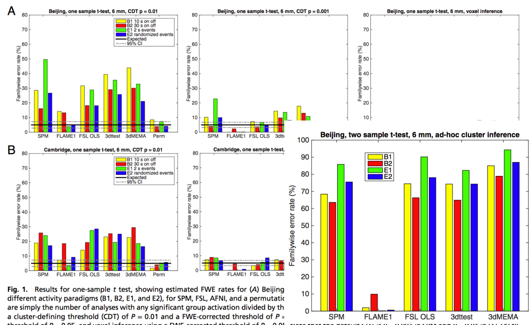


Fig. 1. Results for one-sample *t* test, showing estimated FWE rates for (*A*) Beijing different activity paradigms (B1, B2, E1, and E2), for SPM, FSL, AFNI, and a permutatic are simply the number of analyses with any significant group activation divided by th a cluster-defining threshold (CDT) of $P = 0.01$ and a FWE-corrected threshold of P = threshold of $P = 0.05$, and voxel inference using a FWE-corrected threshold of $P = 0.05$. Note that the default CDT is $P = 0.001$ in SPM and $P = 0.01$ in FSL (AFNI does not have a default setting).

Eklund et al., 2016. PNAS

66

66

Multiple Comparison Correction

12 May 2015, RW Cox, 3dClustSim, level 2 (MINOR), type 5 (MODIFY)
Eliminate edge effects of smoothing by padding and unpadding

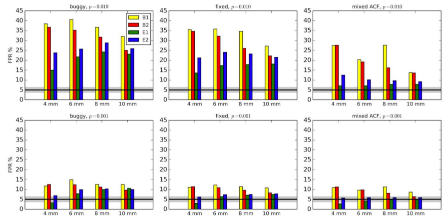
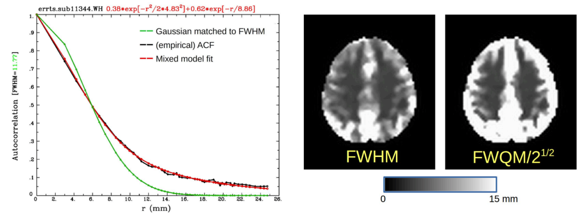


Figure 1. False Positive Rates (FPRs) for various software scenarios, with 1000 2-sample *t*-tests (as in [1,2]) using 20 subjects' data in each sample. "buggy" and "fixed" means the cluster-size thresholds were selected using the Gaussian shape model with the FWHM being the median of the 40 individual subject's values: "buggy" via 3dClustSim before the bug fix, "fixed" via 3dClustSim after the bug fix. "mixed ACF" means the cluster-size threshold was selected using the (Eq 1) model for spatial correlation of the noise, with the *a,b,c* parameters being the median of the 40 individual subject's values (estimated via program 3dFWHMx). Two different per-voxel *p*-value thresholds (1-sided tests, as used in [2]) are shown. The black line shows the nominal 5% false positive rate (out of 1000 trials), and the gray band shows its theoretical 95% confidence interval, 3.65-6.35%. As in [2], different smoothing values were tested (4-10 mm). B1 = 10 s block; B2 = 30 s block; E1 = regular event related; E2 = randomized event related.

Cox et al., 2016. bioRxiv

67

Multiple Comparison Correction



Cox et al., 2016. bioRxiv

68

68

Multiple Comparison Correction

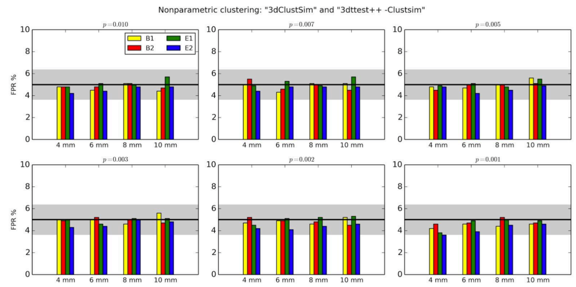


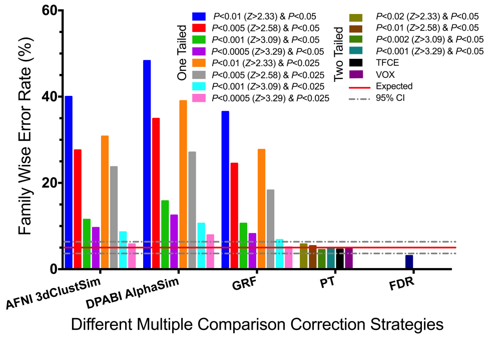
Figure 4. FPRs with cluster-size thresholds now determined from the '-Clustsim' option of 3dttest++ (1-sided tests with 1st nearest neighbor clustering). See Fig. 1 for description of labels, but note that the y-axis range has been changed here for visual clarity.

Cox et al., 2016. bioRxiv

69

69

Family wise Error Rate



20 vs. 20 Permutation 1000 times

Chen, Lu, Yan*, 2018. Human Brain Mapping

70

70

Family wise Error Rate

TABLE I. FWER and cluster size of ALFF (smoothness: 7.94 × 7.31 × 6.86) without GSR under corrections of GRF Theory, AFNI 3dClustSim, and DPABI AlphaSim

(One-tailed twice)		AFNI 3dClustSim		DPABI AlphaSim		GRF	
Voxel threshold	Cluster threshold	FWER	Cluster size	FWER	Cluster size	FWER	Cluster size
$P < 0.01$ ($Z > 2.33$)	$P < 0.05$	40.0%	66.05 ± 0.73	48.3%	60.24 ± 1.68	36.5%	69.35 ± 1.09
$P < 0.005$ ($Z > 2.58$)	$P < 0.05$	27.6%	43.59 ± 0.42	34.9%	39.45 ± 1.13	24.5%	46.70 ± 0.75
$P < 0.001$ ($Z > 3.09$)	$P < 0.05$	11.5%	19.98 ± 0.34	15.8%	18.40 ± 0.61	10.6%	21.29 ± 0.46
$P < 0.0005$ ($Z > 3.29$)	$P < 0.05$	9.6%	14.53 ± 0.25	12.5%	13.93 ± 0.54	8.2%	15.82 ± 0.39
$P < 0.01$ ($Z > 2.33$)	$P < 0.025$	30.8%	74.50 ± 1.14	39.0%	67.72 ± 2.36	27.7%	78.96 ± 1.24
$P < 0.005$ ($Z > 2.58$)	$P < 0.025$	23.7%	47.01 ± 0.59	27.1%	44.48 ± 1.60	18.3%	53.48 ± 0.85
$P < 0.001$ ($Z > 3.09$)	$P < 0.025$	8.6%	22.63 ± 0.25	10.6%	21.00 ± 0.87	6.8%	24.94 ± 0.41
$P < 0.0005$ ($Z > 3.29$)	$P < 0.025$	5.8%	17.33 ± 0.22	7.9%	16.03 ± 0.71	5.1%	18.51 ± 0.50

20 vs. 20 Permutation 1000 times

Chen, Lu, Yan*, 2018. Human Brain Mapping

71

71

Family wise Error Rate

TABLE II. FWER under correction of three kinds of cluster-based correction with the strictest threshold, 6 versions of PT-based correction as well as FDR correction

			FWER										
	Voxel threshold	Cluster threshold	ALFF	fALFF	ReHo	DC	VMHC	ALFF with GSR	fALFF with GSR	ReHo with GSR	DC with GSR	VMHC with GSR	ALFF (8 mm smoothed)
Smoothness (mm, x×y×z)			7.94 × 7.31 × 6.86	7.34 × 7.42 × 7.20	9.36 × 8.72 × 8.39	7.86 × 7.97 × 7.81	6.31 × 6.87 × 6.61	7.99 × 7.31 × 6.84	7.32 × 7.41 × 7.19	9.24 × 8.56 × 8.18	8.06 × 8.16 × 8.09	6.11 × 6.61 × 6.37	11.88 × 11.53 × 11.68
AFNI 3dClustSim (one-tailed)	$P < 0.0005$ ($Z > 3.29$)	$P < 0.025$	5.8%	6.1%	7.3%	8.5%	6.0%	5.3%	6.6%	6.9%	6.8%	6.4%	5.5%
DPABI AlphaSim (one-tailed)			7.9%	8.3%	8.5%	10.2%	9.0%	7.8%	7.7%	7.8%	8.3%	9.6%	6.9%
GRF (one-tailed)			5.1%	5.5%	4.9%	7.4%	5.2%	4.8%	5.9%	5.3%	5.1%	6.4%	4.4%
PT cluster extent correction (two-tailed)	$P < 0.02$ ($Z > 2.33$)	$P < 0.05$	5.8%	3.6%	5.8%	4.6%	5.2%	4.8%	3.9%	5.5%	5.2%	4.3%	5.3%
	$P < 0.01$ ($Z > 2.58$)	$P < 0.05$	5.4%	4.0%	5.7%	4.6%	5.5%	5.3%	3.8%	5.3%	5.0%	4.5%	5.4%
	$P < 0.002$ ($Z > 3.09$)	$P < 0.05$	4.5%	4.1%	5.3%	4.8%	4.2%	4.5%	5.0%	5.1%	4.7%	4.3%	4.4%
	$P < 0.001$ ($Z > 3.29$)	$P < 0.05$	4.8%	4.5%	4.5%	4.9%	3.4%	4.3%	4.8%	5.4%	4.2%	3.9%	4.1%
PT TFCE			4.6%	3.9%	5.7%	5.0%	4.3%	5.3%	4.2%	5.5%	4.7%	4.8%	4.6%
PT VOX			4.9%	4.9%	5.7%	3.9%	4.7%	6.0%	4.5%	5.6%	4.0%	4.6%	3.9%
FDR correction			3.1%	3.4%	4.4%	2.4%	3.9%	4.1%	2.8%	3.6%	2.4%	3.5%	1.6%

The smoothness in the second row is the estimated effective smoothness of the final metric maps feed to statistical analysis, and was different form the applied smoothness (4 mm FWHM) in pre-processing. The effective smoothness was used in 3 versions of cluster-based correction (i.e., GRF theory correction, AFNI 3dClustSim and DPABI AlphaSim)

20 vs. 20 Permutation 1000 times

Chen, Lu, Yan*, 2018. Human Brain Mapping

72

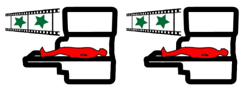
72

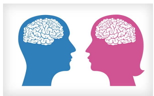
Test-retest Reliability

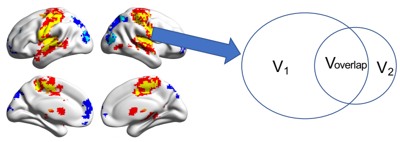
Test-retest reliability

Sex differences in test and retest

Time T1 Time T2







V_1: significant voxels in test

V_2: significant voxels in retest

$V_{overlap}$: voxels significant in both test and retest

Statistical significant voxels

$$Dice = \frac{2 \times V_{overlap}}{V_1 + V_2}$$

Chen, Lu, Yan*, 2018. Human Brain Mapping

73

Test-retest Reliability

TABLE III. Test–retest reliability of sex differences for all R-fMRI metrics with and without GSR under correction of three kinds of cluster-based correction with the strictest threshold, six kinds of PT-based correction and FDR correction, calculated between the first and second sessions in the CORR dataset

			Test–retest reliability (dice coefficient)									
	Voxel threshold	Cluster threshold	ALFF	fALFF	ReHo	DC	VMHC	ALFF with GSR	fALFF with GSR	ReHo with GSR	DC with GSR	VMHC with GSR
AFNI 3dClustSim (one-tailed)	$P < 0.0005$ ($Z > 3.29$)	$P < 0.025$	0.65	0.51	0.50	0.34	0.39	0.64	0.48	0.44	0.28	0.24
DPABI AlphaSim (one-tailed)			0.65	0.51	0.49	0.34	0.39	0.64	0.48	0.45	0.27	0.27
GRF (one-tailed)			0.64	0.51	0.50	0.35	0.39	0.65	0.48	0.43	0.28	0.24
PT cluster extent correction (two-tailed)	$P < 0.02$ ($Z > 2.33$)	$P < 0.05$	0.65	0.70	0.56	0.45	0.40	0.62	0.68	0.45	0.30	0.40
	$P < 0.01$ ($Z > 2.58$)	$P < 0.05$	0.67	0.66	0.52	0.32	0.33	0.60	0.63	0.46	0.27	0.32
	$P < 0.002$ ($Z > 3.09$)	$P < 0.05$	0.63	0.55	0.51	0.36	0.38	0.63	0.52	0.47	0.23	0.32
	$P < 0.001$ ($Z > 3.29$)	$P < 0.05$	0.64	0.51	0.48	0.37	0.38	0.64	0.48	0.44	0.28	0.26
PT TFCE			0.68	0.75	0.54	0.48	0.44	0.66	0.74	0.44	0.31	0.42
~~PT VOX~~			~~0.66~~	~~0.34~~	~~0.48~~	~~0.37~~	~~0.22~~	~~0.65~~	~~0.31~~	~~0.38~~	~~0.11~~	~~0.14~~
FDR correction			0.64	0.67	0.54	0.39	0.37	0.63	0.64	0.47	0.23	0.29

For test–retest reliability for all the 31 kinds of multiple comparison correction strategies, please see Supporting Information Table S13

- Moderate test-retest reliability
- ALFF, fALFF, ReHo are better than DC and VMHC

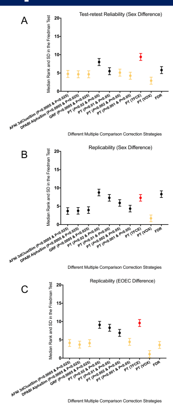
212 M vs. 208 F × 2 times

Chen, Lu, Yan*, 2018. Human Brain Mapping

74

PT with TFCE outperforms

Permutation test TFCE, a strict multiple comparison correction strategy, reached the best balance between family-wise error rate (under 5%) and test-retest reliability / replicability

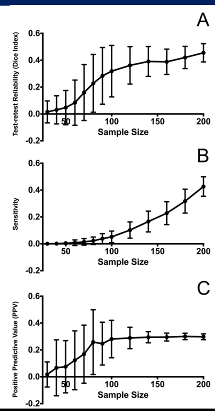


Chen, Lu, Yan*, 2018. Human Brain Mapping

75

Sample Size Matters

Randomly draw k subjects from the "SWU 4" site in the CORR dataset, which has two sessions of 116 males and 105 females



Chen, Lu, Yan*, 2018. Human Brain Mapping

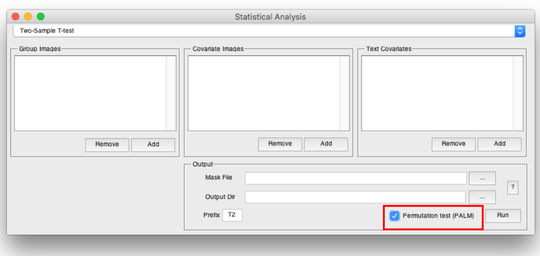
76

Reproducibility of R-fMRI Metrics on the Impact of Different Strategies for Multiple Comparison Correction and Sample Sizes

- Permutation test with TFCE reached the best balance between FWER and reproducibility
- Although R-fMRI indices attained moderate reliabilities, they replicated poorly in distinct datasets (replicability < 0.3 for between-subject sex differences, < 0.5 for within-subject EOEC differences)
- For studies examining effect sizes similar to or even less than those of sex differences, results from a sample size <80 (40 per group) should be considered preliminary, given their low reliability (< 0.23), sensitivity (< 0.02) and PPV (< 0.26).

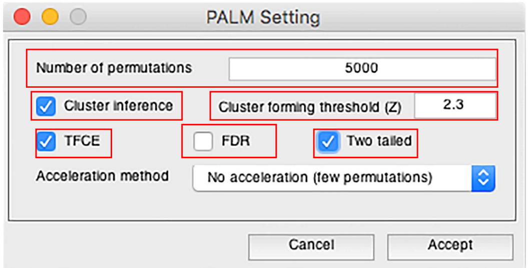
77

Permutation Test



78

Permutation Test

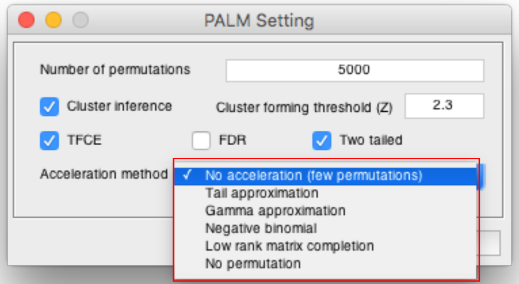


Based on PALM: Winkler, A.M., Ridgway, G.R., Douaud, G., Nichols, T.E., Smith, S.M., 2016. Faster permutation inference in brain imaging. Neuroimage 141, 502-516.

79

79

Permutation Test

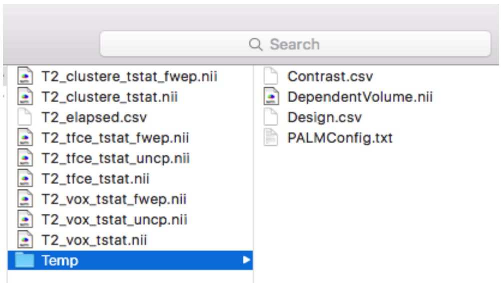


Based on PALM: Winkler, A.M., Ridgway, G.R., Douaud, G., Nichols, T.E., Smith, S.M., 2016. Faster permutation inference in brain imaging. Neuroimage 141, 502-516.

80

80

Permutation Test



Based on PALM: Winkler, A.M., Ridgway, G.R., Douaud, G., Nichols, T.E., Smith, S.M., 2016. Faster permutation inference in brain imaging. Neuroimage 141, 502-516.

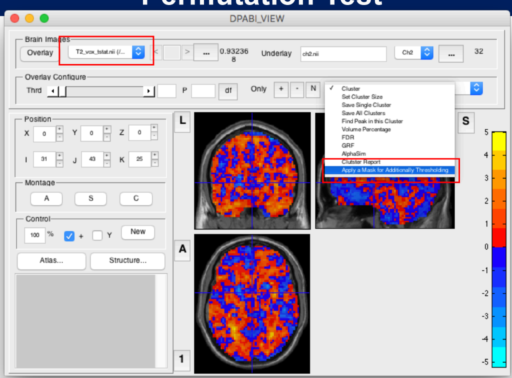
81

Permutation Test

1. _vox_tstat.nii is the T value of a voxel.
2. _vox_tstat_uncp.nii is the p value corresponds to the rank of the observed T value within the permutations FOR A GIVEN VOXEL (the null distribution is the permuted T values of that given voxel). Computing the rank is one of the ways in which the p-value can be obtained (it's then divided by the number of permutations).
3. _vox_tstat_fwep.nii is the p value corresponds to the rank of the observed T value within the permutations of maximum T values across all the voxels (the null distribution is composed by the maximum T value across all the voxels for each permutation). For the corrected, the distribution of the maximum is used as reference, and the rank (or quantile) of a given voxel in relation to that distribution is used to obtain p-values.
3. _clustere_tstat.nii is simply the size (in voxels) of the cluster. This number acts as the test statistic.
4. _clustere_tstat_fwep.nii: p-values computed in the same way as 3, i.e., using the distribution of the maximum cluster size.
5. The TFCE maps are similar to Points 1, 2 and 3.

82

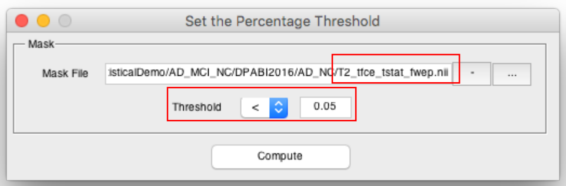
Permutation Test



83

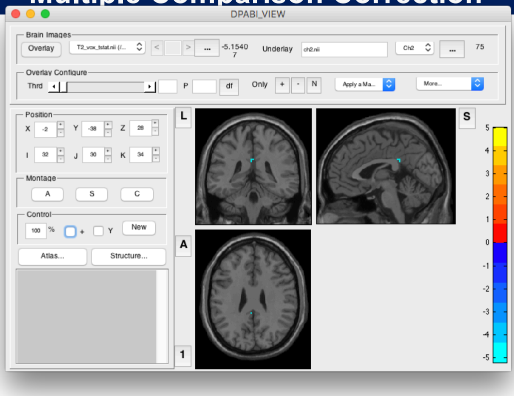
83

Permutation Test



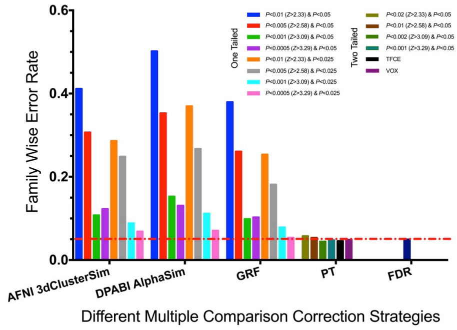
84

Multiple Comparison Correction



85

Multiple Comparison Correction



Chen, Lu, Yan*, Human brain mapp. 2017.

20 vs. 20 Permutation 1000 times

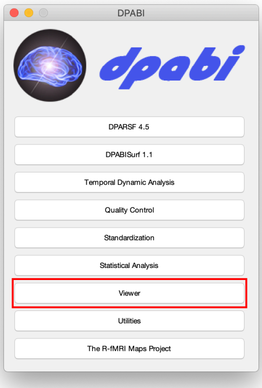
86

Outline

- Quality Control
- Statistical Analysis
- ➡ Results Viewing

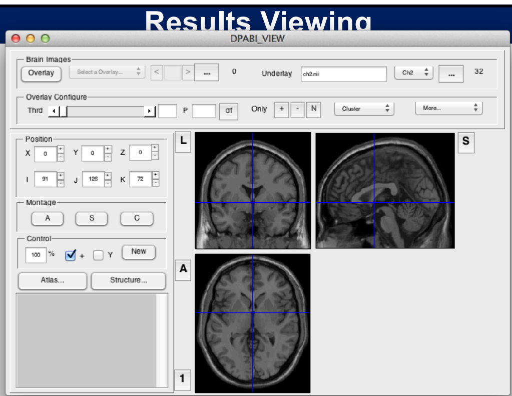
87

87



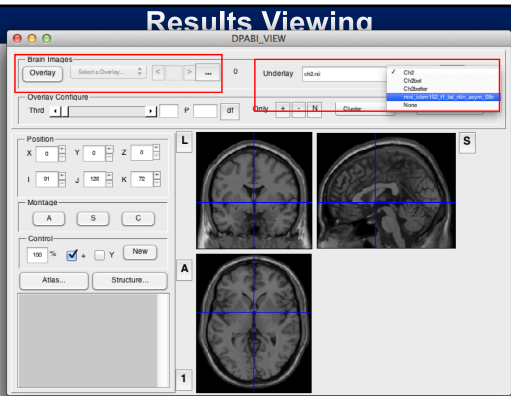
88

Results Viewing



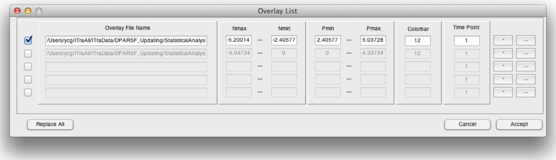
89

Results Viewing



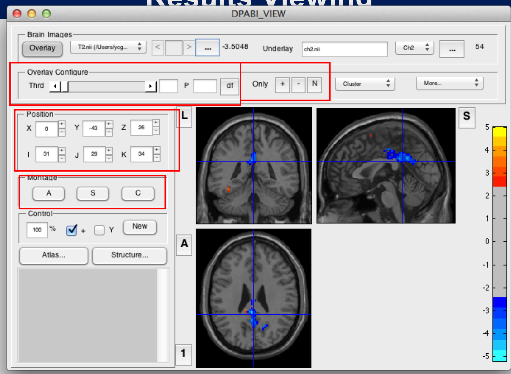
90

Results Viewing



91

Results Viewing



92

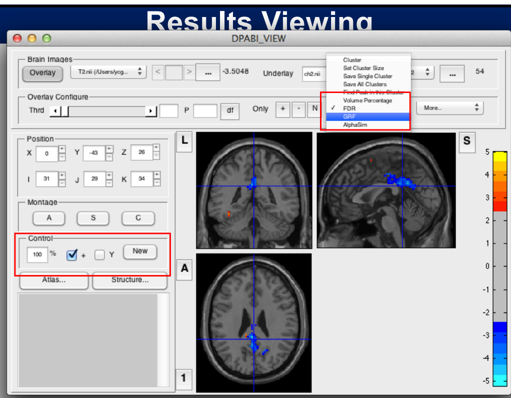
92

Results Viewing

93

93

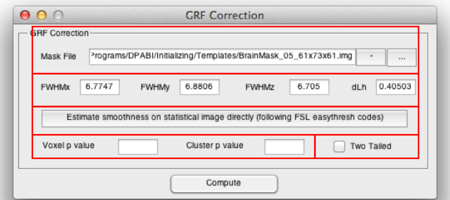
Results Viewing



94

94

Results Viewing

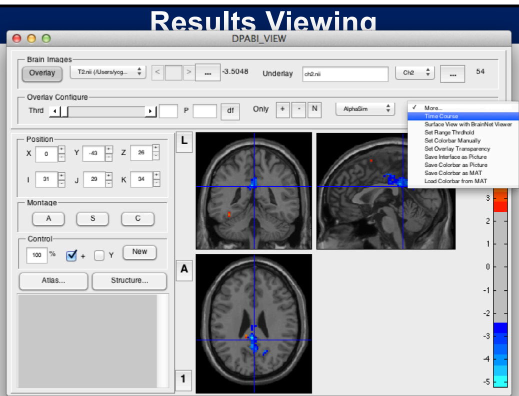


Voxel Z > 2.3, Cluster P < 0.05, Two One-Tailed Corrections:
equivalent to
Voxel P < 0.0214, Cluster P < 0.1, Two Tailed.

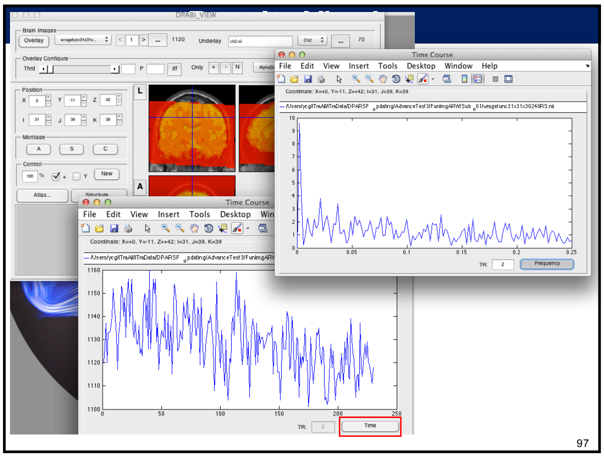
95

95

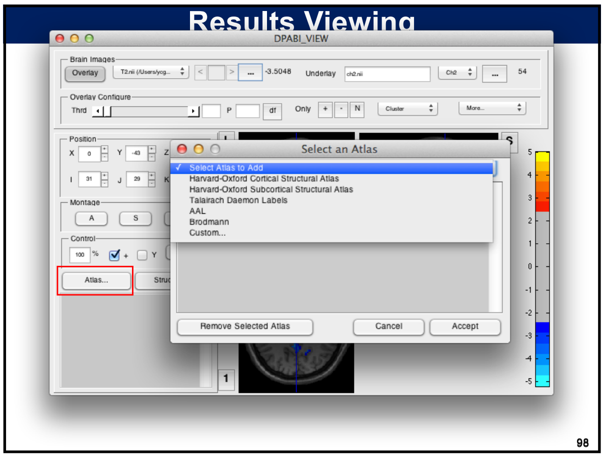
Results Viewing



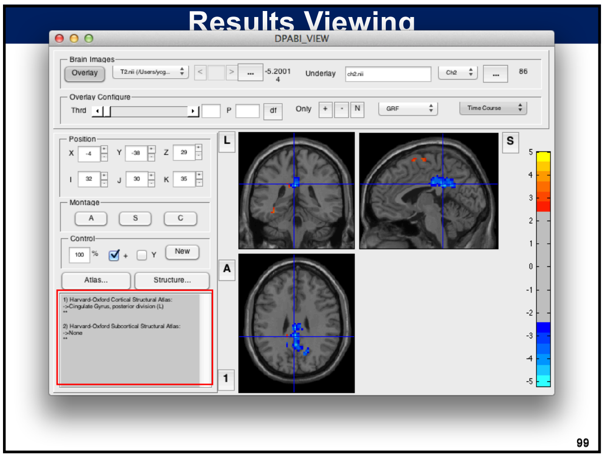
96



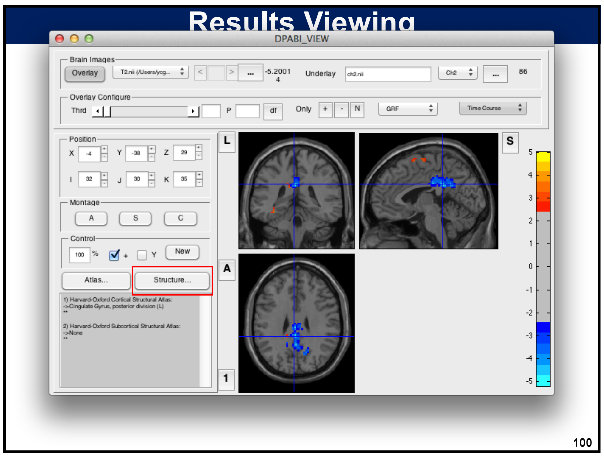
97



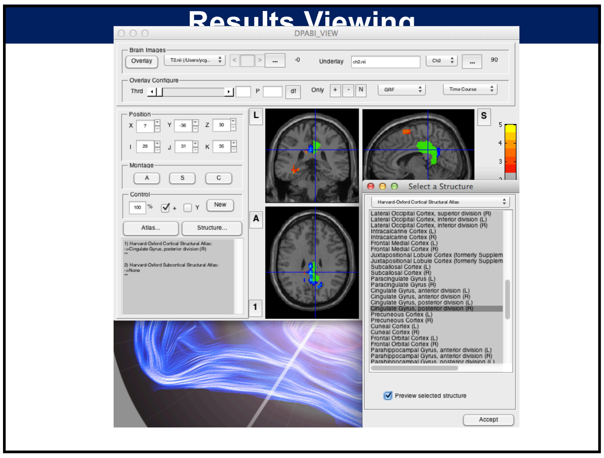
98



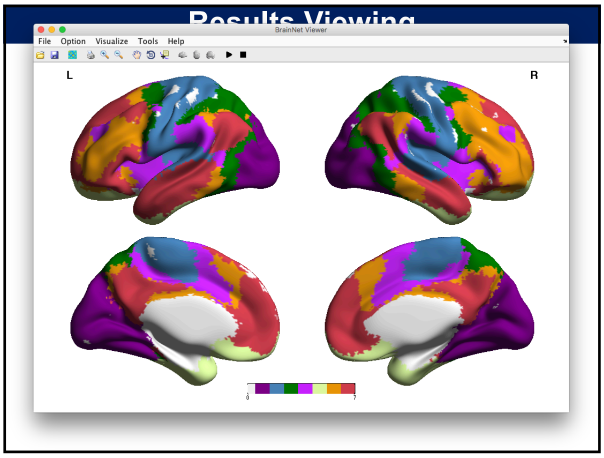
99



100

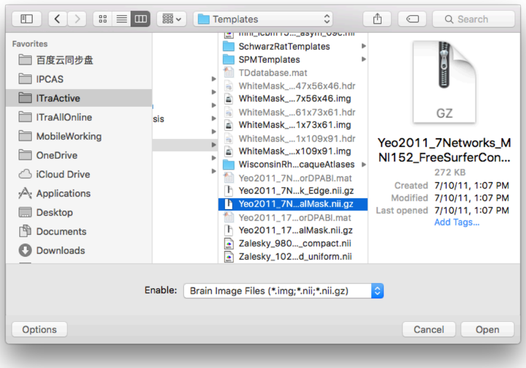


101



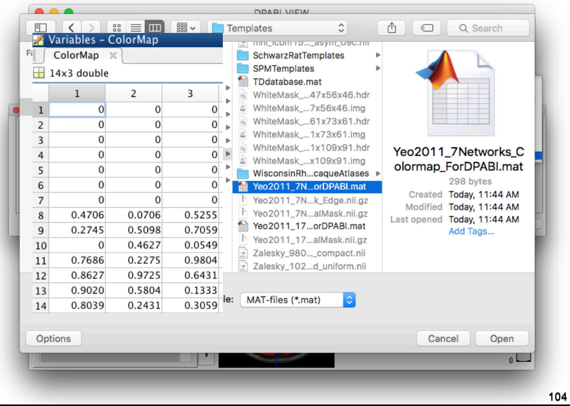
102

Results Viewing



103

Results Viewing

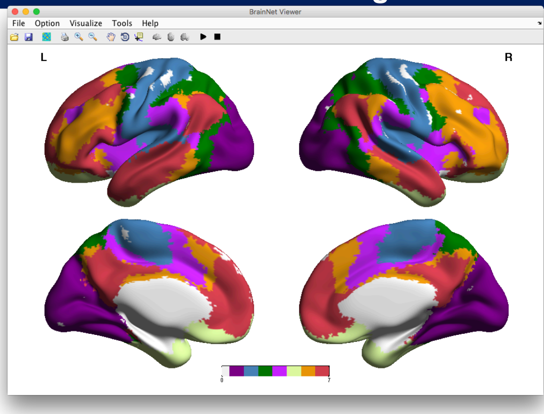


104

Results Viewing

105

Results Viewing

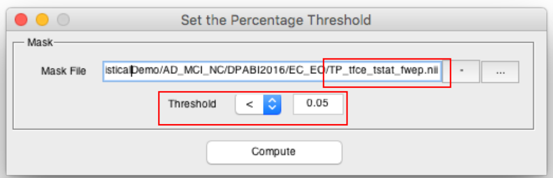


106

Results Viewing

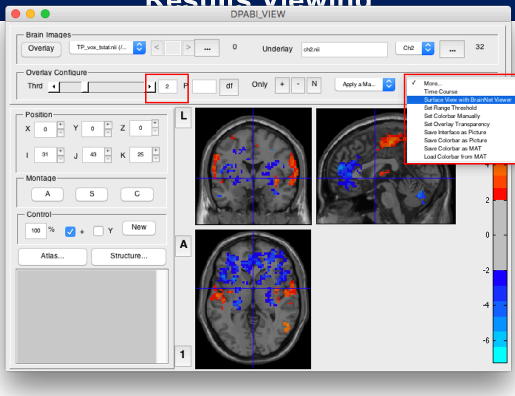
107

Permutation Test



108

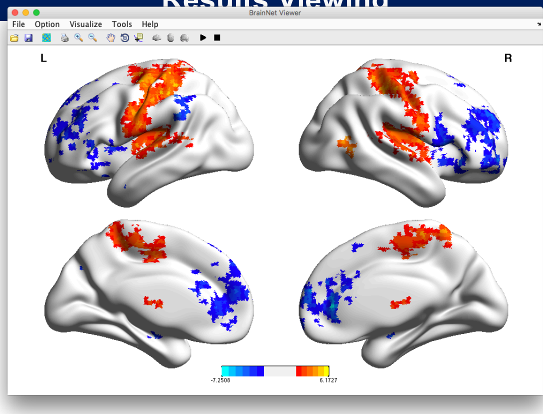
Results Viewing



109

109

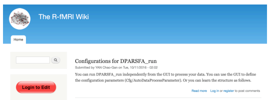
Results Viewing



110

Further Help

http://rfmri.org/Course



http://wiki.rfmri.org

The R-fMRI Journal Club

Official Account: RFMRILab

111

111

Acknowledgments

Chinese Academy of Sciences
Xi-Nian Zuo

Hangzhou Normal University
Yu-Feng Zang

Beijing Normal University
Yong He
Xin-Di Wang

Peking University
Jia-Hong Gao
Wei-Wei Men

NYU Child Study Center
F. Xavier Castellanos

Child Mind Institute
Michael P. Milham

Funding
- National Natural Science Foundation of China
- Chinese Academy of Sciences

112

112

Thanks for your attention!

113

113