Joint analysis of cortical structural measurements through NPC

Anderson M. Winkler

FMRIB Analysis Group

19 October 2016

◆ Human Brain Mapping 37:1486–1511 (2016) ◆

Non-Parametric Combination and Related Permutation Tests for Neuroimaging

Anderson M. Winkler,[1]* Matthew A. Webster,[1] Jonathan C. Brooks,[2] Irene Tracey,[1] Stephen M. Smith,[1] and Thomas E. Nichols[1,3]

[1]*Oxford Centre for Functional MRI of the Brain, University of Oxford, Oxford, United Kingdom*
[2]*Clinical Research and Imaging Centre, University of Bristol, Bristol, United Kingdom*
[3]*Department of Statistics & Warwick Manufacturing Group, University of Warwick, Coventry, United Kingdom*

Abstract: In this work, we show how permutation methods can be applied to combination analyses such as those that include multiple imaging modalities, multiple data acquisitions of the same modality, or simply multiple hypotheses on the same data. Using the well-known definition of union-intersection tests and closed testing procedures, we use synchronized permutations to correct for such multiplicity of tests, allowing flexibility to integrate imaging data with different spatial resolutions, surface and/or volume-based representations of the brain, including non-imaging data. For the problem of joint inference, we propose and evaluate a modification of the recently introduced non-parametric combination (NPC) methodology, such that instead of a two-phase algorithm and large data storage requirements, the inference can be performed in a single phase, with reasonable computational demands. The method compares favorably to classical multivariate tests (such as MANCOVA), even when the latter is assessed using permutations. We also evaluate, in the context of permutation tests, various combining methods that have been proposed in the past decades, and identify those that provide the best control over error rate and power across a range of situations. We show that one of these, the method of Tippett, provides a link between correction for the multiplicity of tests and their combination. Finally, we discuss how the correction can solve certain problems of multiple comparisons in one-way ANOVA designs, and how the combination is distinguished from conjunctions, even though both can be assessed using permutation tests. We also provide a common algorithm that accommodates combination and correction. *Hum Brain Mapp 37:1486–1511, 2016.*

FMRIB
Oxford Centre for Functional MRI of the Brain

Quick GLM review

Model:

$$\mathbf{Y} = \mathbf{M}\boldsymbol{\psi} + \boldsymbol{\epsilon}$$

Null hypothesis:

$$\mathcal{H}_0 : \mathbf{C}'\boldsymbol{\psi} = \mathbf{0}$$

Quick GLM review: Two-sample t test

$$\mathbf{Y} = \mathbf{M}\psi + \epsilon$$

$$\downarrow$$

$$\begin{bmatrix} 0.9670 \\ 0.5472 \\ 0.9727 \\ 0.7148 \\ 0.6977 \\ 0.2161 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} +0.8290 \\ +0.5429 \end{bmatrix} + \begin{bmatrix} +0.1380 \\ -0.2817 \\ +0.1437 \\ +0.1719 \\ +0.1549 \\ -0.3268 \end{bmatrix}$$

$$t = +1.3258$$

How likely is a value at least as large as this if there is no effect?

Quick GLM review: Two-sample t test

1 0	0 1	1 0	1 0	0 1	1 0	1 0	0 1	0 1	1 0
1 0	1 0	0 1	1 0	1 0	0 1	1 0	0 1	1 0	0 1
1 0	1 0	1 0	0 1	1 0	1 0	0 1	1 0	0 1	0 1
0 1	0 1	0 1	0 1	0 1	0 1	0 1	0 1	0 1	0 1
0 1	0 1	0 1	0 1	1 0	1 0	1 0	1 0	1 0	1 0
0 1	1 0	1 0	1 0	0 1	0 1	0 1	1 0	1 0	1 0
$t = +1.33$	$t = -0.91$	$t = +0.25$	$t = -0.93$	$t = +0.42$	$t = +2.24$	$t = +0.41$	$t = -0.45$	$t = -2.40$	$t = -0.47$

0 1	1 0	1 0	0 1	0 1	1 0	0 1	0 1	1 0	1 0
1 0	0 1	1 0	0 1	1 0	0 1	0 1	1 0	0 1	1 0
1 0	1 0	0 1	1 0	0 1	0 1	1 0	0 1	0 1	1 0
1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	0 1
0 1	0 1	0 1	0 1	0 1	0 1	1 0	1 0	1 0	0 1
0 1	0 1	0 1	1 0	1 0	1 0	0 1	0 1	0 1	0 1
$t = +0.47$	$t = +2.40$	$t = +0.45$	$t = -0.41$	$t = -2.24$	$t = -0.42$	$t = +0.93$	$t = -0.25$	$t = +0.91$	$t = -1.33$

Quick GLM review: Two-sample t test

1 0	0 1	1 0	1 0	0 1	1 0	1 0	0 1	0 1	1 0
1 0	1 0	0 1	1 0	1 0	0 1	1 0	0 1	1 0	0 1
1 0	1 0	1 0	0 1	1 0	1 0	0 1	1 0	0 1	0 1
0 1	0 1	0 1	0 1	0 1	0 1	0 1	0 1	0 1	0 1
0 1	0 1	0 1	0 1	1 0	1 0	1 0	1 0	1 0	1 0
0 1	1 0	1 0	1 0	0 1	0 1	0 1	1 0	1 0	1 0
$t = +1.33$	$t = -0.91$	$t = +0.25$	$t = -0.93$	$t = +0.42$	**$t = +2.24$**	$t = +0.41$	$t = -0.45$	$t = -2.40$	$t = -0.47$

0 1	1 0	1 0	0 1	0 1	1 0	0 1	0 1	1 0	1 0
1 0	0 1	1 0	0 1	1 0	0 1	0 1	1 0	0 1	1 0
1 0	1 0	0 1	1 0	0 1	0 1	1 0	0 1	0 1	1 0
1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	1 0	0 1
0 1	0 1	0 1	0 1	0 1	0 1	1 0	1 0	1 0	0 1
0 1	0 1	0 1	1 0	1 0	1 0	0 1	0 1	0 1	0 1
$t = +0.47$	**$t = +2.40$**	$t = +0.45$	$t = -0.41$	$t = -2.24$	$t = -0.42$	$t = +0.93$	$t = -0.25$	$t = +0.91$	$t = -1.33$

Quick GLM review: Two-sample t test

There were 3 cases of a statistic at least as large as the one observed. We have run 20 permutations. Thus:

$$p = \frac{3}{20} = 0.15$$

Quick GLM review

1. Partition the model:

$$\mathbf{Y} = \mathbf{M}\boldsymbol{\psi} + \boldsymbol{\epsilon} \quad \rightarrow \quad \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

2. Choose a permutation strategy (e.g., Freedman–Lane or Dekker).
3. Choose assumptions (EE and/or ISE).
4. Run!

Quick GLM review

Permutation tests are superior:

- Reasonable assumptions: data is exchangeable.
- Wide variety of statistics (but needs pivotality).
- Good for small datasets.
- All information needed is in the dataset itself, not in an idealised population.
- Resilient to outliers.

FMRIB
Oxford Centre for Functional MRI of the Brain

Non-Parametric Combination (NPC)

Non-Parametric Combination (NPC)

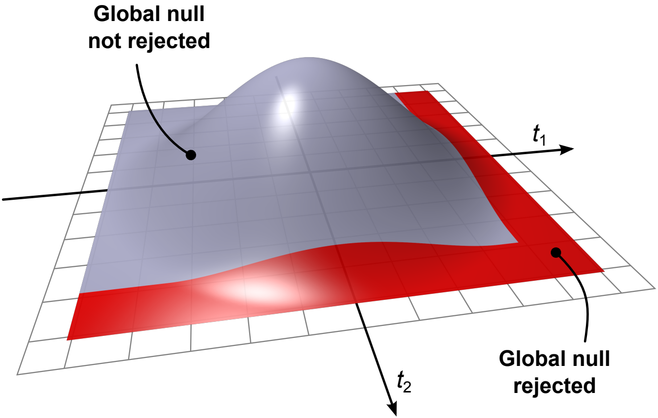
We may conduct multiple tests regarding a certain hypothesis, and none of these may be significant on their own right. However, on the *aggregate*, they may be significant.

Three possibilities:

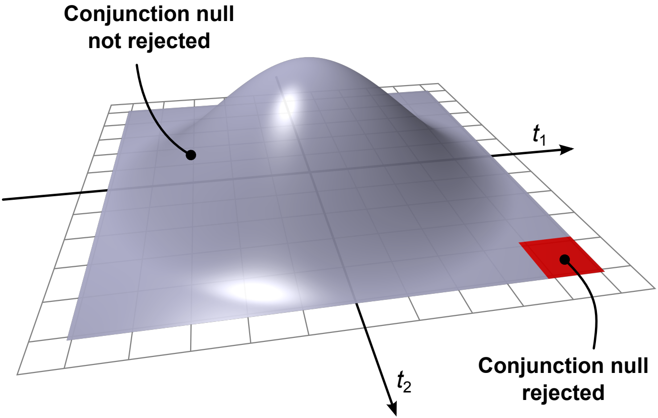
- Reject the null if *any* is significant.
- Reject the null if *all* are significant.
- Reject the null if *an aggregate measure* is significant.

Each individual test is called **partial test**, and used to test a **joint null hypothesis**.

Union-Intersection Test (UIT)



Intersection-Union Test (IUT), i.e., *conjunction test*

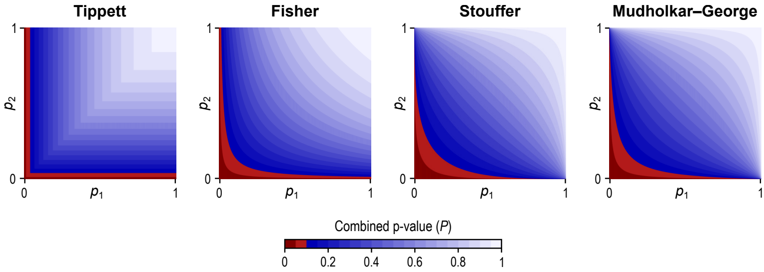


Combining functions

Method	Statistic
Tippett	$\min_k (p_k)$
Fisher	$-2\sum_{k=1}^{K} \ln (p_k)$
Stouffer	$\frac{1}{\sqrt{K}} \sum_{k=1}^{K} \Phi^{-1}(1-p_k)$
Mudholkar–George	$\frac{1}{\pi}\sqrt{\frac{3(5K+4)}{K(5K+2)}} \sum_{k=1}^{K} \ln \left(\frac{1-p_k}{p_k}\right)$

If the tests were guaranteed to be independent, a p-value could be computed using parametric formulas.
Otherwise, use permutations.

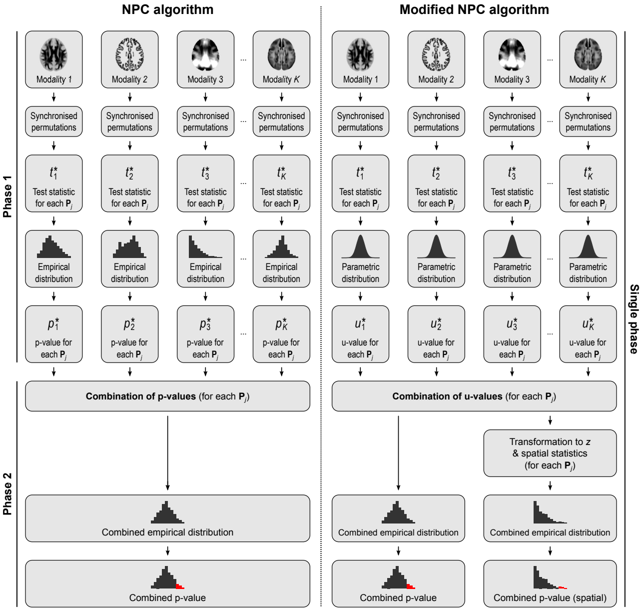
Combining functions



Combining functions (more available)

Method	Test statistic (T)	Significance (p-value, P)
Tippett	$\min_k (p_k)$	$1-(1-T)^K$
Fisher	$-2\sum_{k=1}^{K} \ln(p_k)$	$1-\chi^2(T;\ \nu=2K)$
Stouffer	$\frac{1}{\sqrt{K}}\sum_{k=1}^{K} \Phi^{-1}(1-p_k)$	$1-\Phi(T;\ \mu=0,\ \sigma^2=1)$
Wilkinson	$\sum_{k=1}^{K} I(p_k \leqslant \alpha)$	$\sum_{k=T}^{K} \binom{K}{k}\alpha^k(1-\alpha)^{K-k}$
Good	$\prod_{k=1}^{K} p_k^{w_k}$	$\sum_{k=1}^{K} w_k^{K-1}T^{1/w_k}\left(\prod_{i=1}^{k-1}(w_k-w_i)^{-1}\right)\left(\prod_{i=k+1}^{K}(w_k-w_i)^{-1}\right)$
Lancaster	$\sum_{k=1}^{K} w_k F_k^{-1}(1-p_k)$	$1-G(T)$
Winer	$\sum_{k=1}^{K} t_{\mathrm{cdf}}^{-1}(1-p_k;\ \nu_k) \Big/ \sqrt{\sum_{k=1}^{K}\frac{\nu_k}{\nu_k-2}}$	$1-\Phi(T;\ \mu=0,\ \sigma^2=1)$
Edgington	$\sum_{k=1}^{K} p_k$	$\sum_{j=0}^{\lfloor T\rfloor}(-1)^j\binom{K}{j}\frac{(T-j)^K}{K!}$
Mudholkar–George	$\frac{1}{\pi}\sqrt{\frac{3(5K+4)}{K(5K+2)}}\sum_{k=1}^{K}\ln\left(\frac{1-p_k}{p_k}\right)$	$1-t_{\mathrm{cdf}}(T;\ \nu=5K+4)$
Darlington–Hayes	$\frac{1}{r}\sum_{k=1}^{r}\Phi^{-1}\left(1-p_{(k)}\right)$	Computed through Monte Carlo methods. Tables are available.
Zaykin et al. (TPM)	$\prod_{k=1}^{K} p_k^{I(p_k\leqslant\alpha)}$	$\sum_{k=1}^{K}\binom{K}{k}(1-\alpha)^{K-k}\left(I\left(T>\alpha^k\right)\alpha^k + I\left(T\leqslant\alpha^k\right)T\sum_{j=0}^{k-1}\frac{(k\ln\alpha-\ln T)^j}{j!}\right)$
Dudbridge–Koeleman (RTP)	$\prod_{k=1}^{r} p_{(k)}$	$\binom{K}{r+1}(r+1)\int_0^1(1-t)^{K-r-1}A(T,t,K)\,\mathrm{d}t$
Dudbridge–Koeleman (DTP)	$\max\left(\prod_{k=1}^{r} p_{(k)}, \prod_{k=1}^{K} p_k^{I(p_k\leqslant\alpha)}\right)$	$\sum_{k=1}^{r}\binom{K}{k}(1-\alpha)^{K-k}A(T,\alpha,k)+I(r<K)\binom{K}{r+1}(r+1)\int_0^{\alpha}(1-t)^{K-r-1}A(T,t,K)\,\mathrm{d}t$
Taylor–Tibshirani (TS)	$\frac{1}{K}\sum_{k=1}^{K}\left(1-p_{(k)}\frac{K+1}{k}\right)$	$1-\Phi\left(T;\ \mu=0,\ \sigma^2\approx\frac{1}{K}\right)$
Jiang et al. (TTS)	$\frac{1}{K}\sum_{k=1}^{K} I\left(p_{(k)}\leqslant\alpha\right)\left(1-p_{(k)}\frac{K+1}{k}\right)$	Computed through Monte Carlo methods.

T is the statistic for each method and P its asymptotic significance. All methods are shown as function of the p-values for the partial tests. For certain methods, however, the test statistic for the partial tests, if available, can be used directly. K is the number of tests being combined, p_k, $k=\{1,2,\ldots,K\}$ are the partial p-values, w_k are positive weights assigned to the respective p_k, $p_{(r)}$ are the p_k with rank r in ascending order (most significant first), α is the significance level for the partial tests, $I(\cdot)$ is an indicator function that evaluates as 1 if the condition is satisfied, 0 otherwise, $\lfloor\cdot\rfloor$ represents the floor function, χ^2_ν is the cumulative distribution function (cdf) for a χ^2 distribution, with the ν degrees of freedom, t_{cdf} is the cdf of the Student's t distribution with degrees of freedom ν, and t_{cdf}^{-1} its inverse, Φ is the cdf of the normal distribution with mean μ and variance σ^2, and Φ^{-1} its inverse, and F and G are the cdf of arbitrary, yet well chosen distributions. For the two Dudbridge–Koeleman methods, $A(T,a,b) = I\left(T>a^b\right)a^b + I\left(T\leqslant a^b\right)T\sum_{j=0}^{b-1}(b\ln a - \ln T)^j/j!$.


NPC algorithm
Modified NPC algorithm
Phase 1
Phase 2
Single phase
Modality 1
Modality 2
Modality 3
Modality K
Synchronised permutations
t_1^* Test statistic for each $\mathbf{P}_j$
t_2^* Test statistic for each $\mathbf{P}_j$
t_3^* Test statistic for each $\mathbf{P}_j$
t_K^* Test statistic for each $\mathbf{P}_j$
Empirical distribution
Parametric distribution
p_1^* p-value for each $\mathbf{P}_j$
p_2^* p-value for each $\mathbf{P}_j$
p_3^* p-value for each $\mathbf{P}_j$
p_K^* p-value for each $\mathbf{P}_j$
u_1^* u-value for each $\mathbf{P}_j$
u_2^* u-value for each $\mathbf{P}_j$
u_3^* u-value for each $\mathbf{P}_j$
u_K^* u-value for each $\mathbf{P}_j$
Combination of p-values (for each $\mathbf{P}_j$)
Combination of u-values (for each $\mathbf{P}_j$)
Transformation to z & spatial statistics (for each $\mathbf{P}_j$)
Combined empirical distribution
Combined p-value
Combined p-value (spatial)

Procedure (adapted for imaging)

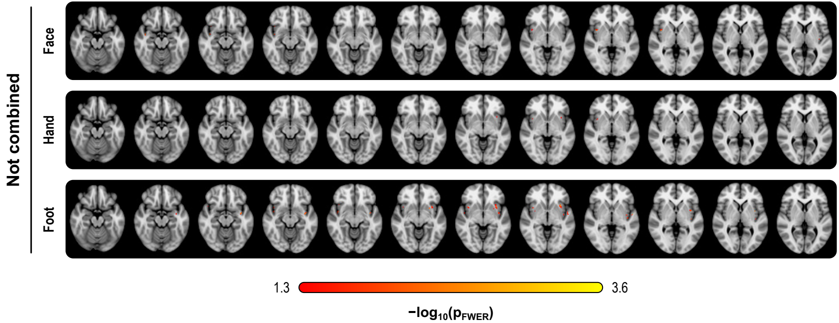
- For each permutation, compute the statistic separately for each modality.
- Convert to a *u*-value (simply a parametric *p*-value with a different name to avoid confusion).
- Combine.
- Repeat many times, and at the end, compute the p-value.

Benefits

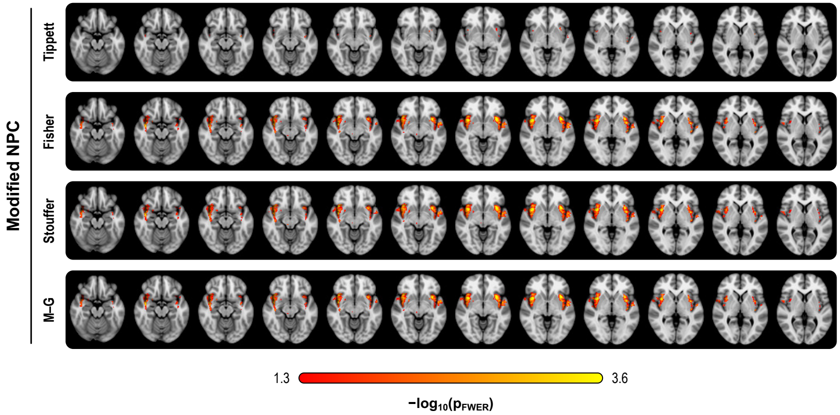
- No need for independence.
- No need to model the non-independence.
- Comes with all other benefits of permutation methods.
- More powerful than MANOVA.

- Needs exchangeability, like any permutation test.

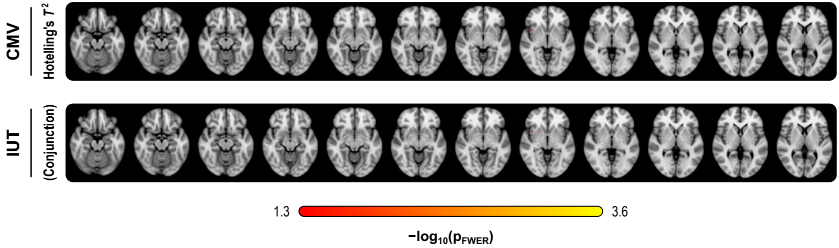
Example: Pain study



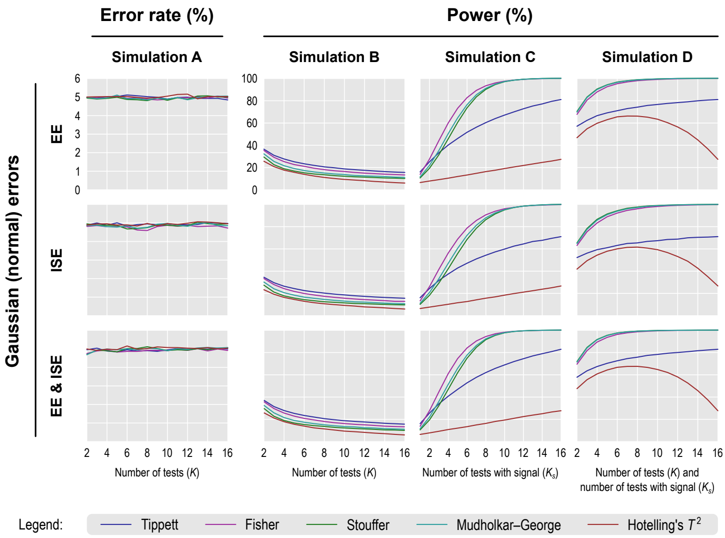
Example: Pain study



Example: Pain study



Error rates and power of different combining functions

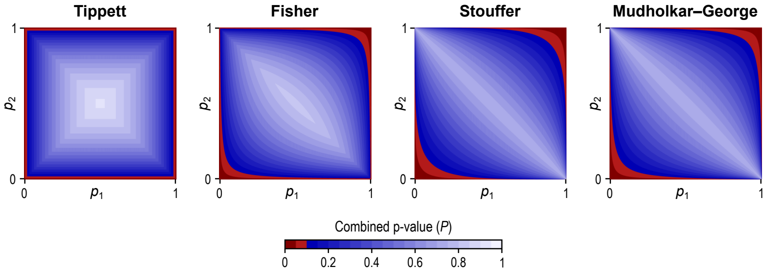


Combination: Concordant directions favoured

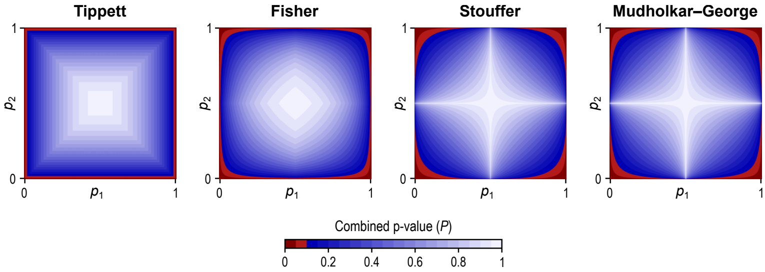
$$T = \max\left(-2\sum_{k=1}^{K}\ln\left(p_k\right), -2\sum_{k=1}^{K}\ln\left(1-p_k\right)\right)$$

Compute the combined statistic, one for each direction, then take the best of both results.

Combination: Concordant directions favoured



Combination: Two-tailed tests (direction irrelevant)



Multiple testing correction

Another type of multiple testing problem:

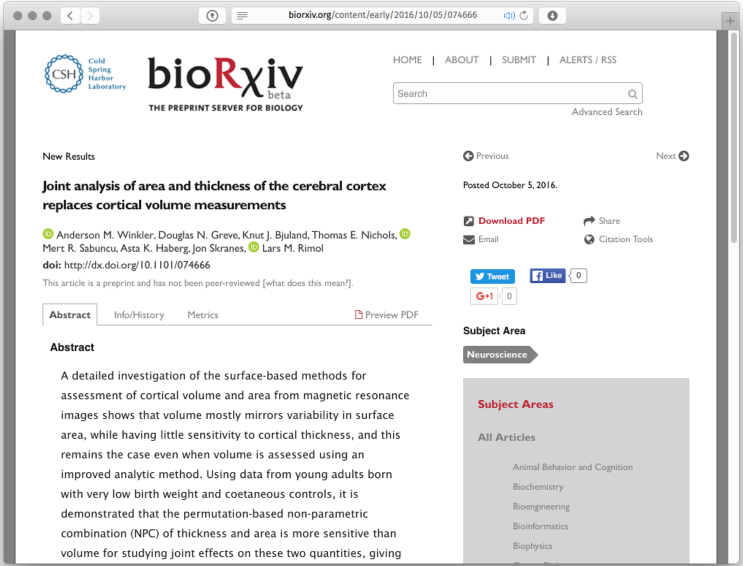
- Multiple hypotheses in the same model (e.g., multiple contrasts).
- Multiple designs (e.g., different seeds).
- Multiple modalities (e.g. multiple DTI measures).
- Multiple pipelines (e.g., different smoothing levels).
- Multiple multivariate hypotheses (e.g. profile analyses).
- Imaging and non-imaging data.

Let's call this *other* multiple testing problem as **MTP-II** to distinguish it from the spatial case, that shall be called **MTP-I**.

Multiple testing correction

Correction over contrasts means you will never have to do an F-test again to account for multiple testing.
It would be dangerous anyway if $\text{rank}(\mathbf{C}) > 2$.

Joint analysis of thickness & area


biorxiv.org/content/early/2016/10/05/074666

Cold Spring Harbor Laboratory

bioRxiv beta

THE PREPRINT SERVER FOR BIOLOGY

HOME | ABOUT | SUBMIT | ALERTS / RSS

Search

Advanced Search

New Results

Previous Next

Posted October 5, 2016.

Joint analysis of area and thickness of the cerebral cortex replaces cortical volume measurements

Anderson M. Winkler, Douglas N. Greve, Knut J. Bjuland, Thomas E. Nichols, Mert R. Sabuncu, Asta K. Haberg, Jon Skranes, Lars M. Rimol

doi: http://dx.doi.org/10.1101/074666

This article is a preprint and has not been peer-reviewed [what does this mean?].

Download PDF Share

Email Citation Tools

Tweet Like 0

G+1 0

Abstract Info/History Metrics Preview PDF

Subject Area

Neuroscience

Abstract

A detailed investigation of the surface-based methods for assessment of cortical volume and area from magnetic resonance images shows that volume mostly mirrors variability in surface area, while having little sensitivity to cortical thickness, and this remains the case even when volume is assessed using an improved analytic method. Using data from young adults born with very low birth weight and coetaneous controls, it is demonstrated that the permutation-based non-parametric combination (NPC) of thickness and area is more sensitive than volume for studying joint effects on these two quantities, giving

Subject Areas

All Articles

Animal Behavior and Cognition

Biochemistry

Bioengineering

Bioinformatics

Biophysics

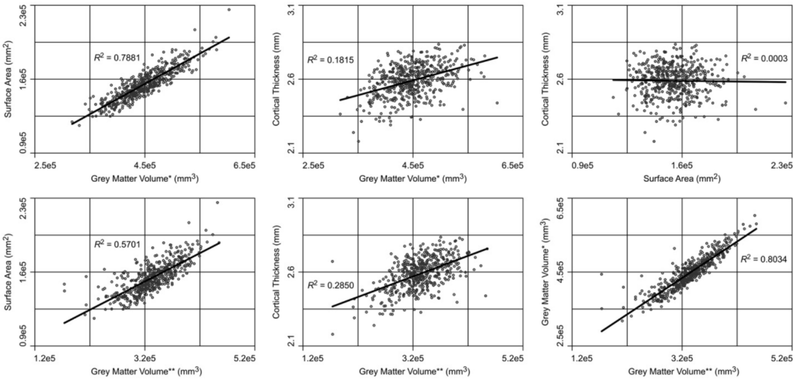
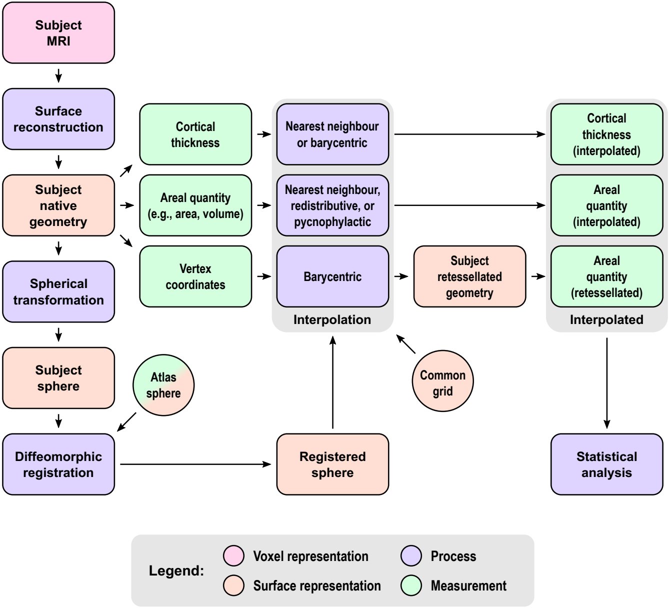
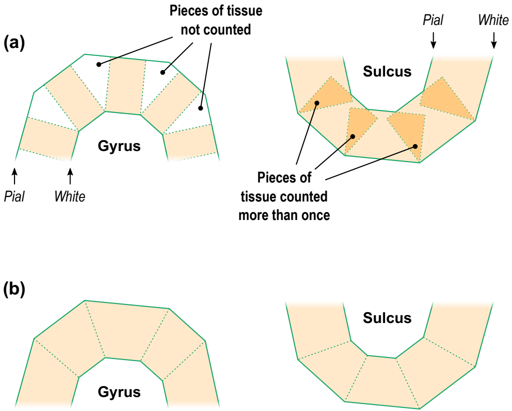
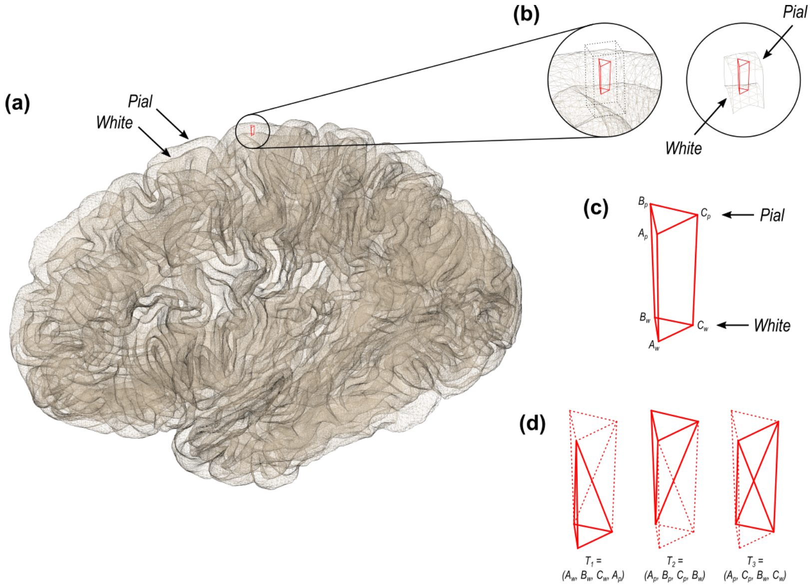


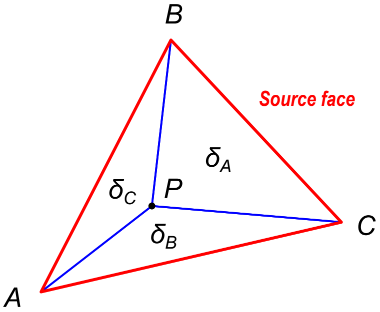
Fig. 3. Correlations between global measurements. Each point represents a pair of measurements for each subject. R^2 is the variance explained by a linear regression model. The significances are shown in Table 2. *Measurement in the surface-based representation. **Measurement in the volume-based representation.


Subject MRI
Surface reconstruction
Subject native geometry
Spherical transformation
Subject sphere
Diffeomorphic registration
Atlas sphere
Cortical thickness
Areal quantity (e.g., area, volume)
Vertex coordinates
Nearest neighbour or barycentric
Nearest neighbour, redistributive, or pycnophylactic
Barycentric
Interpolation
Subject retessellated geometry
Common grid
Registered sphere
Cortical thickness (interpolated)
Areal quantity (interpolated)
Areal quantity (retessellated)
Interpolated
Statistical analysis
Legend:
Voxel representation
Surface representation
Process
Measurement

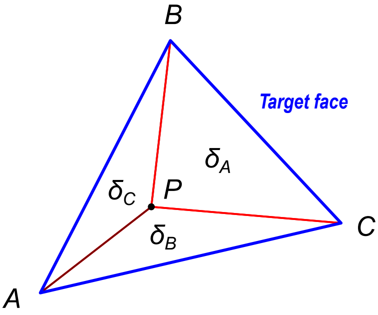

(a)
Pieces of tissue not counted
Pial
White
Sulcus
Gyrus
Pial
White
Pieces of tissue counted more than once
(b)
Sulcus
Gyrus



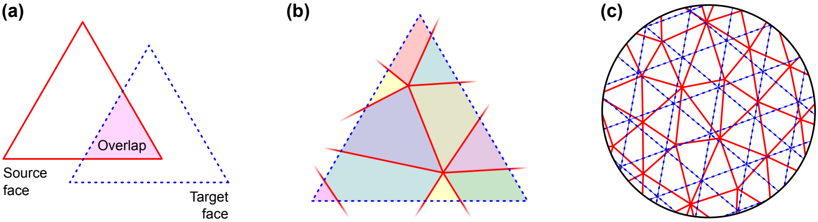
FMRIB Oxford Centre for Functional MRI of the Brain

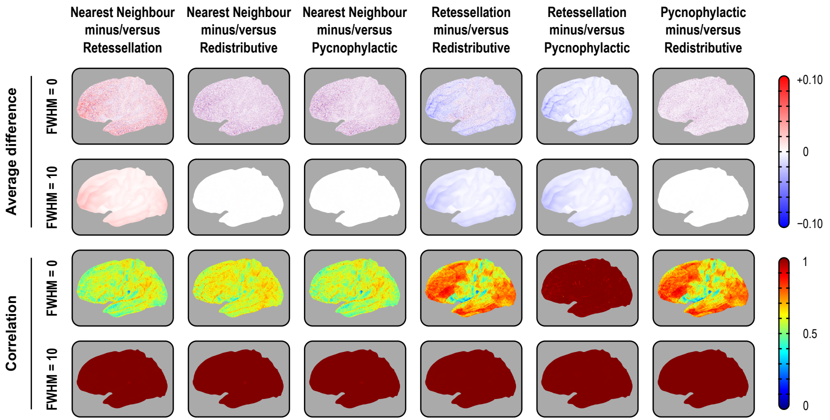


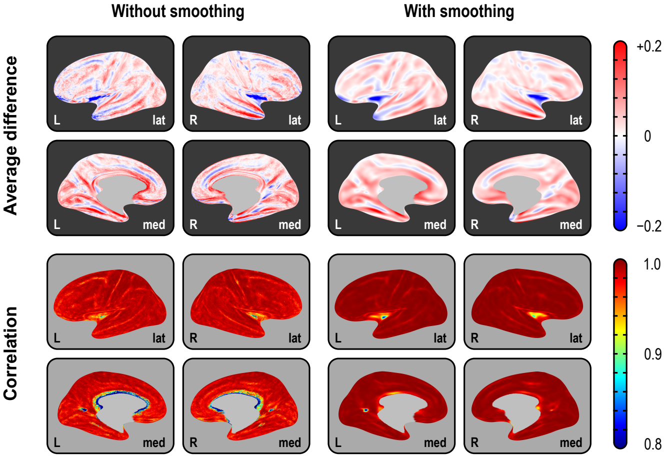
Barycentric

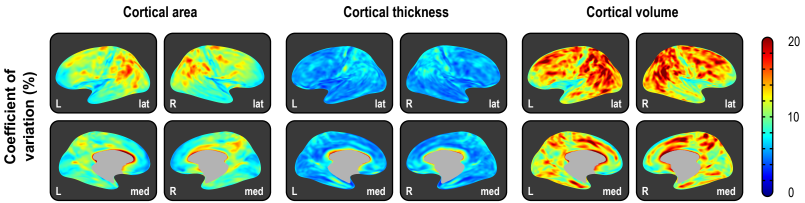


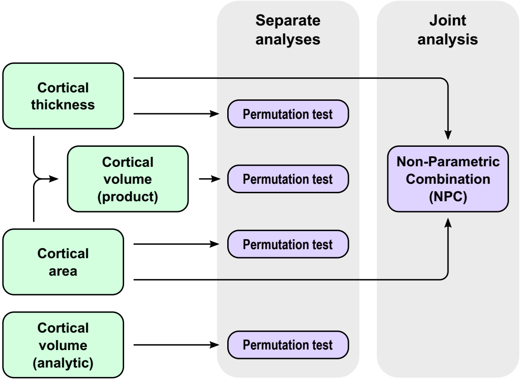
Redistributive

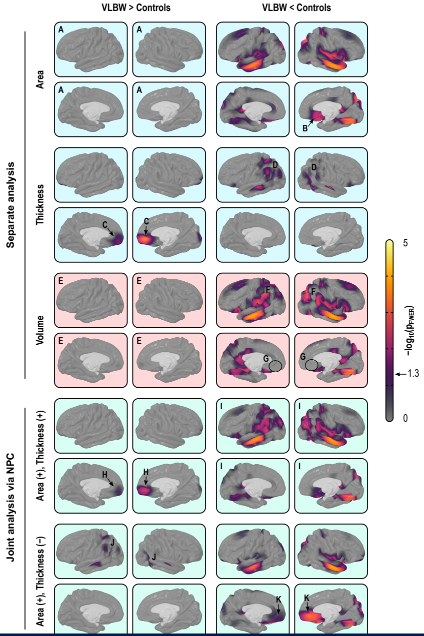

(a)
Overlap
Source face
Target face
(b)
(c)


Nearest Neighbour minus/versus Retessellation
Nearest Neighbour minus/versus Redistributive
Nearest Neighbour minus/versus Pycnophylactic
Retessellation minus/versus Redistributive
Retessellation minus/versus Pycnophylactic
Pycnophylactic minus/versus Redistributive
Average difference
Correlation
FWHM = 0
FWHM = 10
FWHM = 0
FWHM = 10
+0.10
0
−0.10
1
0.5
0


Without smoothing
With smoothing
Average difference
Correlation
L
lat
R
lat
L
med
R
med
+0.2
0
−0.2
1.0
0.9
0.8


Cortical area
Cortical thickness
Cortical volume
Coefficient of variation (%)
L
lat
R
lat
L
med
R
med
20
10
0


Separate analyses
Joint analysis
Cortical thickness
Permutation test
Cortical volume (product)
Permutation test
Non-Parametric Combination (NPC)
Cortical area
Permutation test
Cortical volume (analytic)
Permutation test


VLBW > Controls
VLBW < Controls
Separate analysis
Area
Thickness
Volume
Joint analysis via NPC
Area (+), Thickness (+)
Area (+), Thickness (−)
A
B
C
D
E
F
G
H
I
J
K
5
−log10(pFWER)
1.3
0

Best method in the Universe:

- No mixing in unknown proportions.
- Find effects even if they cancel each other.
- Surface-based.
- Doesn't preclude separate analyses (complementary).
- Non-parametric (even better: permutation-based).
- Born with multiple-testing correction in mind.

```
winkler — bash — 72×26
Last login: Sun May 17 12:56:41 on ttys000
You have mail.
winkler@dyn80:~ $ palm
========================================================================
      ___           ___                         ___
     /  /\         /  /\                       /__/\
    /  /::\       /  /::\                     |  |::\
   /  /:/\:\     /  /:/\:\   ___     ___      |  |:|:\
  /  /:/~/:/    /  /:/~/::\ /__/\   /  /\   __|__|:|\:\
 /__/:/ /:/    /__/:/ /:/\:\\  \:\ /  /:/  /__/::::| \:\
 \  \:\/:/     \  \:\/:/__\/ \  \:\  /:/   \  \:\~~\__\/
  \  \::/       \  \::/       \  \:\/:/     \  \:\
   \  \:\        \  \:\        \  \::/       \  \:\
    \  \:\        \  \:\        \__\/         \  \:\
     \__\/         \__\/                       \__\/

========================================================================
                 Permutation Analysis of Linear Models
========================================================================

The main options are:

-i <file> : Input(s). More than one can be specified, each one preceded
        by its own -i. All input files must contain the same number of
        observations (e.g., the same number of subjects). Except for NPC
        and MV, mixing is allowed (e.g., voxelwise, vertexwise and
```

fsl.fmrib.ox.ac.uk/fsl/fslwiki/PALM

Help Login Search

PALM

Contents

PALM — Permutation Analysis of Linear Models — is a tool that allows inference using permutation methods, offering a number of features not available in other analysis software. These features currently include:

- Ability to work with volumetric and surface-based formats, including facewise data, as well as with non-imaging data;
- A range of various regression and permutation strategies;
- Statistics that are robust to heteroscedasticity;
- Shuffling of sets of observations, to allow, for instance, the analysis of certain designs with repeated measurements, with no missing data;
- Shuffling of observations with complex, tree-like covariance structure (such as for the Human Connectome Project);
- Permutation with sign-flipping (wild bootstrap);
- Modified Non-Parametric Combination (NPC) for joint inference over multiple modalities, or multiple contrasts, or both together, with various combining functions available;
- Classical multivariate statistics (MANOVA, MANCOVA, CCA) for joint inference over multiple modalities, assessed through robust permutation methods, and also parametrically when such approximations exist;
- Correction over multiple contrasts, multiple modalities, for images with or without the same size or geometry, including non-imaging data, controlling the FWER or FDR;
- Various acceleration methods based on the test statistics and their distributions.

PALM requires Matlab or Octave. It can be executed from inside either environment, or directly from the shell. It can also be called from scripts.

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

FMRIB Oxford Centre for Functional MRI of the Brain

How to run

PALM

```
palm -i bh.area -i bh.thickness [...] -npc -o bh.results
```

Split/Merge

```
palm_hemimerge ?h.*
palm_hemisplit bh.results_*
```

That's all folks.