
Local Shape Analysis using MANCOVA

Release 0.00

Beatriz Paniagua, Martin Styner, Marc Macenko, Dimitrios Pantazis and Marc Niethammer

September 3, 2009

The University of North Carolina at Chapel Hill

Abstract

Gross shape measures such as volume have been widely used in statistical analysis of anatomical structures. Statistical shape analysis methods have emerged within the last decade to allow for a localized analysis of shape. Most shape analysis frameworks are though lacking a good statistical underpinning, as they commonly do not allow for the inclusion of independent variables such as age, gender or clinical scores.

This work presents a unified method for local shape analysis that can accomodate different number of variates and contrasts. It also allows to include any number of associated variables in the statistical analysis of the data. Several cases of study are given to clarify the explanation of the different types of data that can be analyzed and the parameters that can be used to tune the program *shapeAnalysisMANCOVA*. This tool has been designed to interact seamlessly with the existing UNC SPHARM-PDM [27] based shape analysis toolbox.

Latest version available at the [Insight Journal](http://hdl.handle.net/1926/1338) [<http://hdl.handle.net/1926/1338>]
Distributed under [Creative Commons Attribution License](#)

Contents

1	Introduction	2
2	Traditional Approach	3
2.1	Local Testing of Group Mean Difference using Hotelling T^2 metric	3
2.2	Correction for Multiple Comparison	4
	Minimum over Non-parametric permutation tests	5
	False Discovery Rate	5
3	Methods	5
3.1	Linear Model	6
3.2	Statistics	7

4	Program contents	8
4.1	Input	8
	Input file	9
	Input arguments	9
4.2	Output	11
	Text files	11
	Mesh files	12
5	Groupwise Analysis Scenarios	12
5.1	Cross-sectional groupwise analysis scenario	12
	Hypothesis formulation	12
	Results	13
5.2	Longitudinal groupwise analysis scenario	15
	Results	15
5.3	Cross-sectional with interaction variables groupwise analysis scenario	16
	Results	18
6	Conclusions	18

1 Introduction

Quantitative shape features such as volumetric measurements have been widely used in statistical shape analysis of diverse anatomical structures in previous years. Volume changes are intuitive features as they might explain morphological variation due to illness, but structural local changes at specific locations are not sufficiently reflected in volume measurements. Shape analysis has thus become of increasing interest due to its potential to precisely locate morphological changes between healthy and pathological structures. One of the first and most influential research in shape analysis was presented by DArcy [28] in his ground-breaking book *On Growth and Form*. In more recent years, several researchers proposed shape analysis via medial shape computing and analysis using M-rep [24]. Medial shape analysis separately analyzes the two medial shape properties: local position and thickness [23]. However, this analysis methods were not corrected for multiple, correlated statistical tests. [19], and Golland [9] proposed shape analysis on medial shape descriptions in 3D and 2D, respectively. They used a fixed topology sampled model with implicit correspondence that is fitted to the objects. Another set of studies can be found on shape analysis via deformable registration to a template ([5], [13], [3], [4]). In this studies, inter-subject comparisons are made by analyzing the individual deformable transformations. This analysis of the transformation fields has to cope with the high dimensionality of the transformation, the template selection problem and the sensitivity to the initial position.

All of these studies are focused in local analysis of specific areas of the structures, via multivariate analysis, but have mostly lacked the ability to include independent clinical variables in that analysis. The work presented in this paper offers statistical shape analysis based on a parametric boundary description (SPHARM [1]) as the point-based model computing method. The point-based models will be analyzed with the methods here proposed using multivariate analysis of covariance (MANCOVA). Here, the number of variates being tested is the dimensionality of our observations. Each point of these observations is a three dimensional displacement vector from the mean. The number of contrasts is the number of equations involved

in the null-hypothesis. In order to encompass varying numbers of variates and contrasts, and to account for independent variables, a matrix computation is performed. This matrix represents the multidimensional aspects of the correlation significance and it can be transformed into a scalar measure by manipulation of its eigenvalues. The various ways of combining these eigenvalues is discussed in section 3.2. Section 3.1 describes the linear model framework used in the analysis, while section 3.2 describe the statistical methods applied. Section 4 describes *shapeAnalysisMANCOVA* usage as well as the input and outputs associated. Finally, sections 5 and 6 present several analysis scenarios in detail and the results and conclusions obtained, respectively.

2 Traditional Approach

After correspondence establishment using SPHARM-PDM [1] [27], alignment and scaling normalization in a shape population, the traditional statistical analysis approach consisted in testing for differences between groups at every surface location. The steps performed in order to obtain the statistical analysis are shown in figure 1.

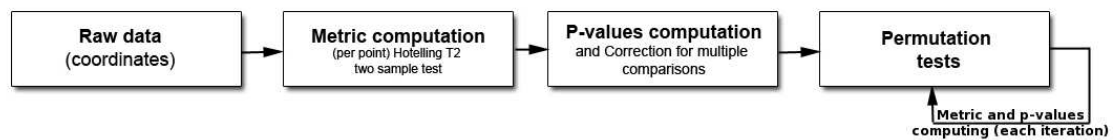


Figure 1: Workflow used in the previous used approach for statistical analysis.

In the following sections each one of these steps will be seen in detail.

2.1 Local Testing of Group Mean Difference using Hotelling T^2 metric

In the previous statistical framework, it was possible to test for differences between two shape populations in two main fashions:

- *Analyzing the magnitude of the local difference vector to a template:* For this option, a template needs to be first selected, usually this template is the common mean of the two groups or the mean of a separate control group. The magnitude of the difference is easily understood and results in difference maps for each subject. Also, the resulting univariate statistical analysis is quite well known and local significance can be easily computed using the Student t metric. The main disadvantage of this method is the need to select a template, which introduces an additional bias into the statistical analysis. We applied this method in earlier studies of hippocampi [26] and ventricles [25].
- *Analyzing the spatial location of each point:* For this option, no template is necessary and multivariate statistics of the (x, y, z) location is necessary. We have chosen to use the Hotelling T^2 two sample difference metric as a measurement of how two groups locally differ from each other. The standard Hotelling T^2 is defined as $T^2 = (\mu_1 - \mu_2)'(\Sigma(\frac{1}{n_1} + \frac{1}{n_2}))^{-1}(\mu_1 - \mu_2)$, where $\Sigma = (\Sigma_1(n_1 - 1) + \Sigma_2(n_2 - 1))/(n_1 + n_2 - 2)$ is the pooled covariance matrix. An alternative modified Hotelling T^2 metric is less sensitive to group differences of the covariance matrixes and the number of samples[21]: $T^2 = (\mu_1 - \mu_2)'(\Sigma_1\frac{1}{n_1} + \Sigma_2\frac{1}{n_2})^{-1}(\mu_1 - \mu_2)$. All our current studies are based on this modified Hotelling T^2 metric.

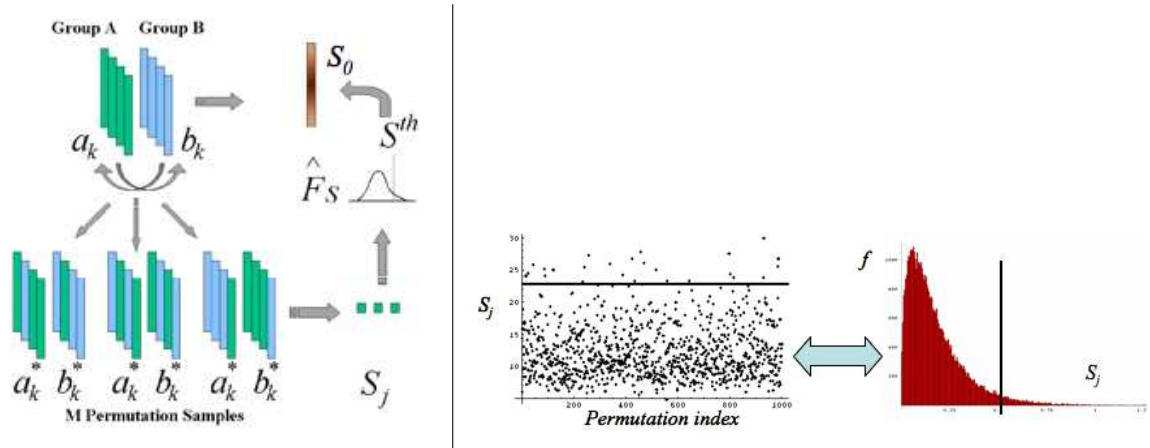


Figure 2: Left: Scheme of P-value computation via permutation tests. The real group difference S_0 (e.g. Hotelling T^2) is compared to that the group differences S_j of random permutations of the group labels. The quantile in the S_j histogram associated with S_0 is the p-value. Right: Example run on test data with 1000 random permutations. On the left the individual S_j are plotted and on the left the corresponding S_j histogram is shown. The black line indicates the value of S_0 .

The significance of the mean difference metric (univariate: Student t, multivariate: Hotelling T^2) was next evaluate at each spatial location via a permutation test approach. Permutation tests are a valid and tractable approach for such an application, as they rely on minimal assumptions and can be applied even when the assumptions of the parametric approach are untenable. Non-parametric permutation tests are exact, distribution free and adaptive to underlying correlation patterns in the data. Further, they are conceptually straightforward and, with recent improvements in computing power, are computationally tractable.

The null hypothesis is that the distribution of the locations at each spatial element is the same for every subject regardless of the group. Permutations among the two groups satisfy the exchangeability condition, i.e. they leave the distribution of the statistic of interest unaltered under the null hypothesis. Given n_1 members of the first group $a_k, k = 1 \dots n_1$ and n_2 members of the second group $b_k, k = 1 \dots n_2$, we can create $M \leq \binom{n_1 + n_2}{n_2}$ permutation samples. A value of M from 20000 and up should yield results that are negligibly different from using all permutations [6].

2.2 Correction for Multiple Comparison

The local shape analysis involves testing from a few to many thousands of hypothesis (one per surface element) for statistically significant effects. It is thus important to control for the multiple testing problem, and the most common measure of multiple false positives is the familywise error rate (FWER). The multiple testing problem has been an active area of research in the functional neuroimaging community. The most widely used methods in the analysis of neuroimaging data use random field theory [2] [29] and make inferences based on the maximum distribution. In this framework, a closed form approximation for the tail of the maximum distribution is available, based on the expected value of the Euler characteristic of the thresholded image [29]. However, random field theory relies on many assumptions such as the same parametric distribution at each spatial location, a point spread function with two derivatives at the origin, sufficient smoothness to justify the application of the continuous random field theory, and a sufficiently high threshold for the asymptotic results to be accurate.

In our work, we are employing non-parametric permutation tests and false discovery rate as two alternative correction methods for the multiple comparison problem.

Minimum over Non-parametric permutation tests

Non-parametric methods deal with the multiple comparisons problem [17] and can be applied when the assumptions of the parametric approach are untenable. Similar methods have been applied in a wide range of functional imaging applications [14, 18, 22].

Details of this method are described in [17] and only a summary is provided here. The correction method is based on computing first the local p-values using permutation tests. The minimum of these p-values across the surface is then computed for every permutation (see Figure 2 left). The appropriate corrected p-value at level α can then be obtained by computing the value at the α -quantile in the histogram of these minimum values. Using the minimum statistic of the p-values, this method correctly controls for the FWER, or the false positives, but no control of the false negatives is provided. The resulting corrected local significance values can thus be regarded as pessimistic estimates akin to a simple Bonferroni correction.

False Discovery Rate

Additionally to the non-parametric permutation correction, a False Discovery Rate Estimation (FDR) [30, 7] method was also implemented and applied. This procedure controls the expected proportion of false positives only among those tests for which a local significance has been detected. The FDR method thus allows an expected proportion (usually 5%) of the FDR corrected significance values to be falsely positive. The correction using FDR provides an interpretable and adaptive criterion with higher power than the non-parametric permutation tests. FDR is further simple to implement and computationally efficient even for large datasets.

The FDR correction (see Figure 2 right) was computed as follows [7]:

1. Select the desired FDR bound q , e.g. 5%. This is the maximum proportion of false positives among the significant tests that you are willing to tolerate (on average).
2. Sort the p-values smallest to largest.
3. Let p_q be the p-value for the largest index i of the sorted p-values $p_{\text{sort},i} \leq q \cdot i/N$, where N is the number of vertices.
4. Declare all locations with a p-value $p \leq p_q$ significant.

3 Methods

The statistical framework proposed in this document differs from the traditional statistical analysis approach (see Figure 1) mainly in the inclusion of a step that consists in computing a General Linear Model (GLM) to test group differences at every surface location. The steps performed in this new proposed approach are shown in figure 3.

In the following sections each one of these steps will be seen in detail. Some examples will be used in order to clarify the text.

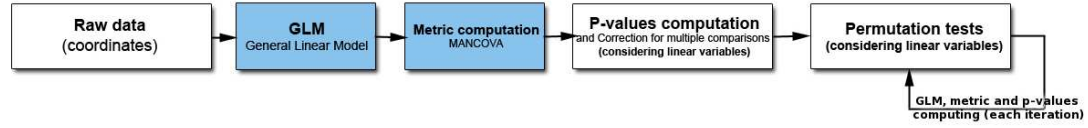


Figure 3: Workflow used for statistical analysis.

3.1 Linear Model

The main innovation in the new approach proposed is that we use a GLM to fit the observed raw data as follows:

$$Y = XB + U \quad (1)$$

Where Y ($n \times d$) is the observation matrix, X ($n \times p$) the design matrix, B ($p \times d$) is the parameter matrix and U ($n \times d$) the error matrix. n = number of subjects, d = dimensionality of features, p = depends on design matrix. B is unknown, X and Y are known.

The parameter matrix B can be computing by solving the equation system as:

$$\hat{B} = (XX^T)^{-} X^T Y; (p \times d) \quad (2)$$

Where X^{-} represents the pseudoinverse of the matrix.

Once parameter matrix B is known, the residual information after applying the GLM to the data is estimated as:

$$E = (Y - X\hat{B})^T (Y - X\hat{B}); (p \times p) \quad (3)$$

This residual matrix E indicates the main eigenvalues of the data represented in the GLM.

EXAMPLE: A case of study in which it is necessary to analyze shape changes of the caudate nuclei in an autistic population compared with a normal population, the observation matrix Y contains row-wise the location vector (x, y, z) ($d = 3$) for each subject both autistic and controls. X , the design matrix, contains the information of each subject variable. Nominal subject variables such as gender and diagnosis group are encoded as a set of binary, mutually exclusive variables. In case of considering gender, this would lead to a 4 binary variable b_0 contains 1 if the subject is male and control, and 0 otherwise, b_1 contains 1 if the subject is female and control and 0 otherwise, b_2 contains 1 if the subject is male and autistic patient and 0 otherwise, b_3 contains 1 if the subject is female and autistic patient and 0 otherwise.

$$Y = \begin{bmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{bmatrix}$$

$$X = \begin{bmatrix} b_0^1 & b_1^1 & b_2^1 & b_3^1 & age^1 & IQ^1 & score^1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_0^n & b_1^n & b_2^n & b_3^n & age^n & IQ^n & score^n \end{bmatrix}$$

\hat{B} will contain the estimates for the contributions relative to gender μ_i , age g , IQ q and scores c . All of these estimates are d-dimensional (3 in this example).

$$\hat{B} = \begin{bmatrix} \mu_0^x & \mu_0^y & \mu_0^z \\ \mu_1^x & \mu_1^y & \mu_1^z \\ \mu_2^x & \mu_2^y & \mu_2^z \\ \mu_3^x & \mu_3^y & \mu_3^z \\ g_1^x & g_1^y & g_1^z \\ q_1^x & q_1^y & q_1^z \\ c_1^x & c_1^y & c_1^z \end{bmatrix}$$

3.2 Statistics

Two matrices A and C are designed to capture the null hypothesis, satisfying the relation:

$$AB = C \quad (4)$$

In the example, using the B matrix given in section 3.1, the null hypothesis to test would be that there is no difference between female autistic subjects and female normal controls and there is no difference between the male autistic subjects and the male normal controls. This hypothesis would be formulated with:

$$A = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The embedding of the null hypothesis into the matrices as described in equation (4) is very simple. When the multiplication AB is made, it computes the difference between the average coordinates of the normal control and the autistic subjects. The top row is for males and the bottom row is for females. By having $C = 0$, we are hypothesizing that there is no difference between the two groups or: $H_0 : \mu_{NC} = \mu_{Autistic}$.

Having the remainder of the A matrix as zeros allows for these variables to be ‘corrected for’. This is accomplished because by including these variables in the linear model that calculates B , their affect is modulated out, leaving a more accurate estimate of the average location of the points. We then formulate the null hypothesis with only these locations (the shape), and not the other factors that we are not currently evaluating.

In order to compute significance, we first calculate the MANCOVA hypothesis matrix:

$$H = (A\hat{B} - C)^T [A(X^T X)^{-1} A^T]^{-1} (A\hat{B} - C) \quad (5)$$

We now compute the three eigenvalues λ_i of HE^{-1} and we are given a choice of the following tests on these eigenvalues from [20]:

$$Wilks : \Lambda = \Pi(1 + \lambda_i)^{-1} \quad (6)$$

$$Hotelling : T_g^2 = (n - r) \sum \lambda_i \quad (7)$$

$$Pillai : V^{(s)} = \sum \frac{\lambda_i}{1 + \lambda_i} \quad (8)$$

$$Roy : \lambda_{max} = \max(\lambda_i) \quad (9)$$

For Wilks Λ we employ a left-sided permutation test (min) in order to compute the p-value, whereas for Hotellings T_g^2 , Pillais $V^{(s)}$ and Roys λ_{max} we compute a right-sided permutation test (max) for the p-value. Any of these statistics can be used in the analysis by specifying the associated arguments.

4 Program contents

The novel tool developed, *shapeAnalysisMANCOVA*, uses CMake [16] as its cross-platform builder. This allows build scripts such as makefiles or project files to be customized to the specific machine.

In order to compile *shapeAnalysisMANCOVA*, ITK, VTK, GenerateCLP and Boost Libraries for C (v 0.39.1) must already be installed. CMake will ask for the directory where the aforementioned library files are located before creating the build scripts.

4.1 Input

The input arguments to the program are mainly for controlling the testing and output preferences, and to point to the data to be analyzed. The input file name is input in the first place in the command line and it will contain the list of mesh files that we want to analyze, as well as scale factors, categorical and associated variables. Combining the arguments input to the program and the information contained in the input file, two main tests can be performed:

- *Group Tests* If a simple group difference test is desired, it is necessary to specify the group being tested. By default it is the first group defined in the data input file, but you can choose which group to test by specifying the right argument ('-testColumn 1'). The default significance level for the p-values computing is 0.05, but any other value can be given on the command line as '-significanceLevel 0.01'. The number of permutations allowed in the statistical significance computing can be also changed. The default value of 10,000 is heuristically sufficient. More advanced analysis of the confidence interval for the p-value can be found in [21]. This value can be set higher (for higher confidence in the results) or lower (for increased speed) with the argument '-numPerms 5000'.
- *Interaction Tests* If corrected statistics using independent variable(s) are desired, it will be also necessary to specify the variable tested. In 4, the independent variable given is the age of the child at the time of the scan. The program will ignore this data unless the argument relative to independent variables is specified. When specifying the proper argument, the program performs the Interaction Tests that will correct the statistics computed using the specified independent variables. That is to say, *shapeAnalysisMANCOVA* correlates any independent variable with the shape at each point when performing the statistical computing. This is done with the permutation test methods described in 1. Spearman and Pearson correlation coefficients are also calculated if needed, testing the correlation

of the independent variable with the shape changes. These shape changes will be quantified in two different ways: first by the length of the mean difference vector at each location and the projection of that vector onto the outward pointing normal on the mean surface. Both of these values are signed positive or negative depending on the direction of the vector of the mean difference vector compared to the surface normal vector.

Both input arguments and input file will be described in detail in the next sections.

Input file

Following the example used in section 3.1 an input file can be composed to perform the statistical analysis. In Figure 4, the first column contains a 0 for autistic or a 4 for normal control (diagnostic group type). In the same way, the second column offers information regarding to gender group type. This is additional data that will set up two groups, diagnostic group and gender group, and it needs to be specified ('-numGroupTypes 2' argument). Depending on the number of group types N, the system will compute all combinations among them i.e. 2^N combinations. The third column in figure 4 contains the scale factor for each mesh file. Each point in the mesh associated with this subject will be scaled by this amount, but only if the correspondent command line argument is also present ('-scale'). The next column is an absolute or relative path to the mesh file. An error will be generated if the file does not contain a mesh or does not exist. The last column (or columns) contains data regarding to the linear variables associated with the mesh. This data will be used in the creation of the GLM and in case of performing Interaction Tests.

0	1	1.126454	/data/caud/5168-003-01_caudL_pp.meta	2.193018
0	0	1.137706	/data/caud/5171-003-01_caudL_pp.meta	2.469541
0	0	0.951810	/data/caud/5174-003-01_caudL_pp.meta	2.759754
0	0	0.861418	/data/caud/5176-004-01_caudL_pp.meta	2.751540
0	0	1.013828	/data/caud/5177-004-01_caudL_pp.meta	2.349076
4	0	1.186670	/data/caud/5052-004-01_caudL_pp.meta	2.387406
4	1	0.989758	/data/caud/5057-003-01_caudL_pp.meta	1.902806
4	0	1.244891	/data/caud/5063-005-01_caudL_pp.meta	2.149213
4	0	1.071684	/data/caud/5082-004-01_caudL_pp.meta	2.012320
4	1	0.964605	/data/caud/5128-004-01_caudL_pp.meta	2.245038
4	1	1.071813	/data/caud/5140-004-01_caudL_pp.meta	2.160164
4	0	1.152571	/data/caud/5142-005-01_caudL_pp.meta	2.918549
4	0	0.987110	/data/caud/5149-003-01_caudL_pp.meta	2.220937

Figure 4: A screenshot of an example mesh-list input file.

Input arguments

shapeAnalysisMANCOVA offers different statistical functionalities that can allow the user to perform and design a statistical analysis. The input arguments allow to tune the program in order to obtain the desired results.

- '**<std::string>**' (required) File which contains all of the data to be analyzed. File is a list pointing to other data sources.
- '**--processinformationaddress <std::string>**' Address of a structure to store process information (progress, abort, etc.). (default: 0)

- `'--xml'` Produce xml description of command line arguments (default: 0)
- `'--echo'` Echo the command line arguments (default: 0)
- `'--trendCorrelation'` For the simple (non-MANCOVA) correlation, test for correlation. (default: 0)
- `'--positiveCorrelation'` For the simple (non-MANCOVA) correlation, test for positive correlation. (Only works in conjunction with parametric testing!) (default: 0)
- `'--negativeCorrelation'` For the simple (non-MANCOVA) correlation, test for negative correlation. (Only works in conjunction with parametric testing!) (default: 0)
- `'--pillai'` Uses Pillai statistic for MANCOVA testing. (default: 0)
- `'--hotelling'` Uses Hotelling statistic for MANCOVA testing. (default: 0)
- `'--wilks'` Uses Wilks statistic for MANCOVA testing. (default: 0)
- `'--roy'` Uses Roy statistic for MANCOVA testing. (default: 0)
- `'--debug'` Outputs additional debugging information. (default: 0)
- `'--scale'` Scale all point data by the cube-root of the scale factor found in the input file. This is typically used for correcting differences such as varying intercranial volumes (ICV). (default: 0)
- `'--simpleCorrsParaP'` Computes the p-value parametrically for the simple correlation test. (default: 0)
- `'--simpleCorrs'` Simple Spearman and Pearson correlations are computed, based on the normal to the average shape. This option is only valid in interaction test mode. (default: 0)
- `'--interactionTest'` Instead of a group test, simply test for statistically significant interaction of the data with a single independent variable. (default: 0)
- `'--computeScaleFactorFromVolumes'` Reinterprets the scaling column values as volumes and compute the scaling factor from them. WARNING: This is different from the traditional file format where these scaling were already pre-computed. (default: 0)
- `'--writeZScores'` Writes out the z-scores. Only uses group assignments from the first group column currently. z-score of group A is computed with mean and standard deviation from group B and vice versa. z-scores are output based on the projections on the mean surface normal as well as the corresponding Mahalanobis distances (where no projection is performed). (default: 0)
- `'--significanceLevel <double>'` What cutoff of p-values is considered significant. Only affects the FDR corrected results. (default: 0.05)
- `'--testColumn <int>'` Which column of the data is being tested (0 based). If this is a standard 'Group Test', it will be one of the group columns. If it is an 'Interaction Test', it will be one of the independent variable columns. (default: 0)
- `'--numIndependent <int>'` Number of different independent variables associated with each subject. (default: 0)
- `'--numGroupTypes <int>'` Number of different classification types. (default: 1)

- `'--numPerms <int>'` Number of permutations to perform. (default: 10000)
- `'--out <std::string>'` All output files will have this as their base. (default: statResult)
- `'--ignore_rest'` Ignores the rest of the labeled arguments following this flag.
- `'--version'` Displays version information and exits.
- `'--help'` Displays usage information and exits.

4.2 Output

shapeAnalysisMANCOVA outputs diverse files. The first type of output files are given as text files containing a header and a list of values associated with each point in the surface, those list of values can be one dimensional (scalar) or three-dimensional (vectors). The header contains the type of data, the dimensionality of data, and the number of points included. The second type of output files contain full shape data in form of mesh definitions. These are composed of a list of point locations, and then a list with full connectivity information for the triangles connecting each point. All these files all have a `'*.meta'` extension and are in ITK Spatial object mesh format. They can be easily visualized using *KWMeshVisu* [15] and 3D Slicer (v3.3 and above). Both types of files will be described next.

Text files

- `'_diffMesh.txt'` includes the difference vectors obtained from the subtraction of two mean shapes from the groups tested.
- `'_normDistance.txt'` includes the magnitude of the difference vectors.
- `'_normProjections.txt'` includes the magnitude of the projection of the previously described difference vectors to its correspondent normal vector in the mean population surface. The magnitude of this projection is a measure of shape variation that can be useful in the analysis.
- `'_mancovarawP.txt'` includes the raw p-values or associated statistical significance for each point for the test performed.
- `'_mancovaFDRP.txt'` includes the p-values after correction for false discovery. False Discovery Rate (FDR) [8] threshold is calculated first, and then the raw p-values are scaled considering that the FDR threshold is now at the given significance level.
- `'_mancovaBonferroniP.txt'` includes the p-values after Bonferroni correction. Bonferroni correction (FDR) threshold is calculated first, and then the raw p-values are scaled considering that the Bonferroni threshold is now at the given significance level.
- `'*_Spearman.txt'` includes a collection of files are generated when the right argument (`'-simpleCorrs'`) is used in an Interaction Test. Those files will include raw P values maps, corrected FDR maps, Bonferroni corrected P values maps, and the Spearman correlation coefficients maps for a population of shapes correlated with a certain independent variable. The correlation is calculated for both normDistProjections shape scalar maps and normProjections shape scalar maps.

- ‘*_Pearson.txt’ includes a collection of files where correlation is used in an Interaction Test. In the same way as the one described for Spearman correlation, the files will contain raw P values maps, corrected FDR maps, Bonferroni corrected P values maps, and the Pearson correlation coefficients maps for a population of shapes correlated with a certain independent variable. The correlation is calculated for both normDistances shape scalar maps and normProjections shape scalar maps. Both Pearson and Spearman correlation related files are always computed.
- ‘*_localZ.txt’ includes the Z-scores maps for a given population of shapes in a Group Test. This score is a dimensionless quantity derived for each point in a surface from group A by subtracting the group B mean from the individual raw shape score and then dividing the difference by the group B standard deviation. This score indicates how much each individual in group A point varies from the composite model of all the surfaces in group B. For subjects in group B, this score is computed with respect to the composite model of all surfaces in group A.

Mesh files

- ‘_meanAll_uncorrected.meta’ includes the overall population mean shape without GLM correction.
- ‘_meanA.meta’ and ‘_meanB.meta’ contain mean shape of each one of the two groups being tested.
- ‘_GLM_corrected_meanAll.meta’ is calculated only in case of Interaction Tests, and it contains the overall mean shape corrected using the values of the indicated independent variables. The General Linear Model (GLM) used in this correction follows equation (1).

5 Groupwise Analysis Scenarios

In order to further clarify the methods and to show *shapeAnalysisMANCOVA* statistical power, three analysis scenarios are going to be exposed in the following sections. The first analysis scenario shows the results and conclusions obtained from a cross-sectional groupwise study. This scenario is the same one used as example in section 3. The second analysis scenario presents the results obtained in a longitudinal study, trying to capture special features of a certain disease progression in different subjects, that will be also compared in a groupwise fashion. Finally, a third analysis scenario will show the power of the tool to correlate shape variation groupwise features with clinical scores.

5.1 Cross-sectional groupwise analysis scenario

Data from 52 autistic and 22 typically developing children between 18 and 35 months (2 years) of age was studied to investigate shape changes associated with autism (see also [11]). The caudate nucleus in each brain scan were expertly segmented and preprocessed as discussed in [27].

Hypothesis formulation

Depending on the hypothesis formulated for any specific group test, the permutation is performed on individual sets of parameters of the design matrix X . The necessary parameters for the Group Test design will be input in the final program as different arguments, that will be discussed in detail in section 4.1.

$$X = \begin{bmatrix} b_0^1 & b_1^1 & b_2^1 & b_3^1 & age^1 & IQ^1 & score^1 \\ & & & \dots & & & \\ b_0^n & b_1^n & b_2^n & b_3^n & age^n & IQ^n & score^n \end{bmatrix}$$

Several hypothesis could be formulated in order to perform different Group Tests in this specific case:

Hypothesis 1: Is there a significant difference between groups after correction for gender, age, IQ and clinical behavioral score? Permute binary variables of group identification (keeping the gender constant). Create a permutation that leads to switches of b_0 and b_2 , as well as b_1 and b_3 .

Example command: 'infile - -numGroupTypes 2 - -numIndependent 3'

Hypothesis 2: Is there a significant age influence after correction for group, gender, IQ and clinical score? Simply permute the age parameter in X .

Example command: 'infile - -numGroupTypes 2 - -numIndependent 3 - -interactionTest'

Hypothesis 3: Is there a significant clinical score influence after correction for group, gender, age and IQ? Simply permute the clinical score parameter in X .

Example command: 'infile - -numGroupTypes 2 - -numIndependent 3 - -interactionTest - -testColumn 2'

Results

All the color maps calculated will be displayed on the average surfaces of the mesh population. In case of performing a Group Test using the diagnostic group (autistic vs normal controls) the mean difference vectors are show in figure 5 and figure 6 (left). Those vectors show the differences in the morphology of the average surface of autistic caudate nuclei surfaces compared with the morphology of the average normal control surface. These mean difference vectors show clearly an overall bending action on the autistic average caudate, displaying vectors pointing away from the head and tail.

The magnitude of the projection of these difference vectors onto the normals colour map is shown in 6 (right). This map shows shape group changes over the average surface. White color is used where there are no relevant changes (magnitudes equals zero). The color gets darker as the magnitudes of change increase. The distance is positive if the mean surface of the autistic population is protruding outside of the mean surface of normal controls; the distance is negative if the mean surface of autistic population is shrinking below the mean surface of normal controls. Positive distances are color-coded by blue and negative distances are color-coded by red.

Figure 7 (left) shows statistical significance for morphological changes. A color coded p-values map can be visualized over a mean surface. In the map, highly significant correlations ($p < 0.01$) are color-coded with red, intermediately significant correlations are color-coded with green ($0.01 > p > 0.05$) and non-significant correlations are color-coded with blue. The main area of significance for our case of study is located on the head of the caudate. This area of the caudate has been demonstrated to be highly involved in learning and memory and also associated with repetitive tasks performance.

Behavioral scores were used to perform Interaction Tests. Morphological changes analysis correlated with the overall behavioral score results are shown in figure 7 (right).

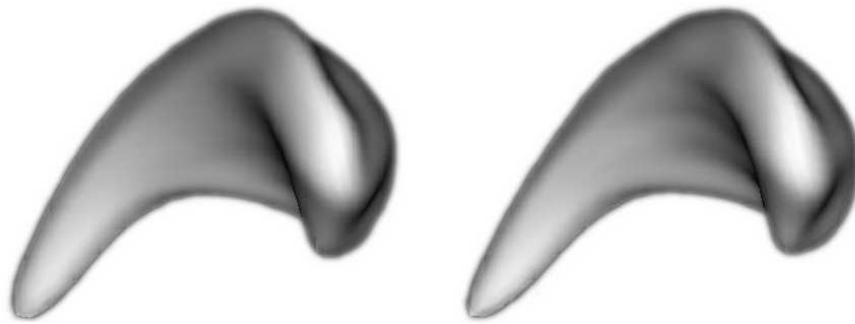


Figure 5: Left: Mean shape of the caudate for the autistic population. Right: Mean shape of the caudate for the normal control (NC) population.

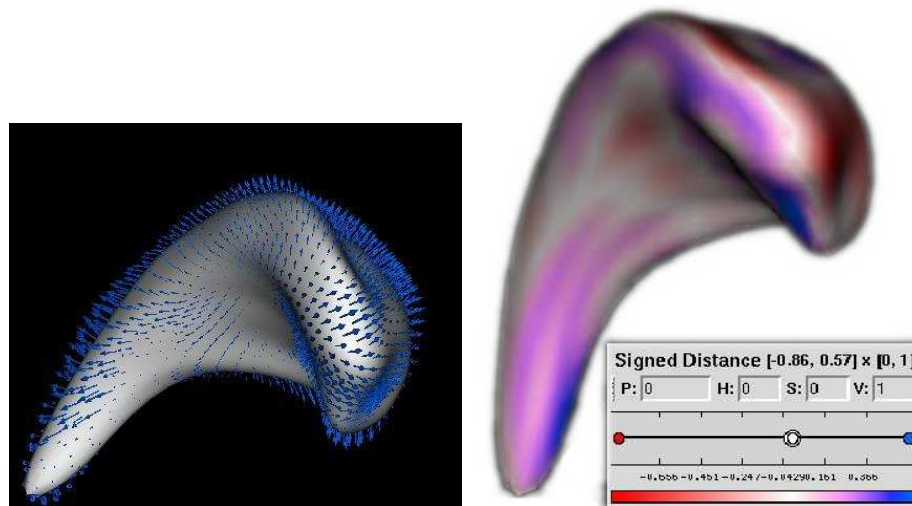


Figure 6: Left: Mean shape of the caudate for the NC population showing the vectors that point to the autistic mean shape. Right: The value of the mean difference vector projected onto the signed normal. Notice that the top of the head is moving negatively (into the shape) while the bottom of the head and the tail are moving positively (out of the shape). This seems to indicate a bending motion.

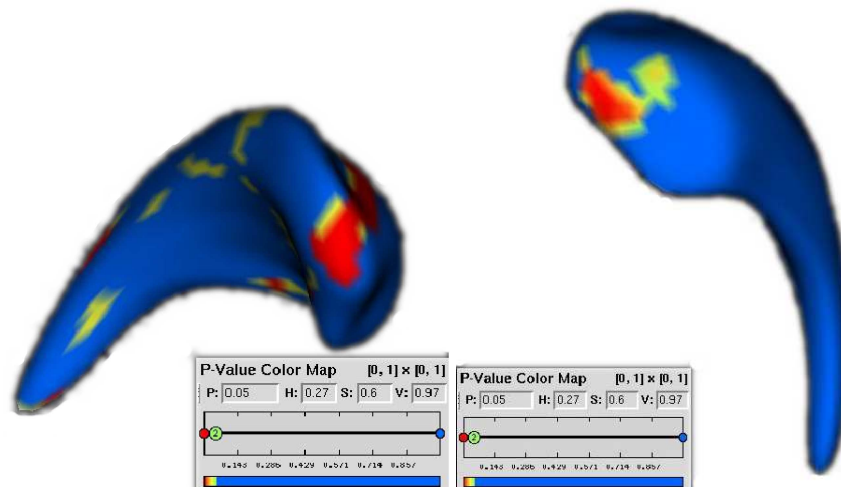


Figure 7: Left: The raw p-values show that there is significant local shape difference mainly in the head of the caudate. Right: The raw p-values show that there is significant local shape correlation with the overall behavioral score, mainly in the head of the caudate.

5.2 Longitudinal groupwise analysis scenario

In addition to groupwise analysis, *shapeAnalysisMANCOVA* also allows to test for differences between the surfaces of two different time points screenings in case of a same patient. The results displayed in this section are part of a schizophrenia study, and tries to find differences among subjects that progressed to schizophrenia and subjects that did not progress from the prodromal state using analysis of covariance. This analytic approach was used first for paired longitudinal comparisons between the baseline measurement and the subsequent follow-up measurements of the same patient. Three features were obtained from that analysis, all them based in the shape differences between the baseline and the follow up status of each subject: direction and magnitude of shape changes, as well as the associated statistical significance. The next step consisted in finding differences between groups of the previously computed features, in a similar fashion than in 5.

Results

Visualization of the results obtained from the aforementioned tests (figures 8 - 10) provide three kinds of information at a specific location on the surface:

- Average magnitude and direction of the shape changes between progressers and non-progressers for schizophrenia: White color is used where the magnitudes are zero. The color gets darker as the magnitudes of the distance increase. The distance is positive if the mean surface of group A (progressers) is protruding outside of the mean surface of group B (non-progressers); the distance is negative if the mean surface of group A is shrinking below the mean surface of group B. Positive distances are color-coded by blue and negative distances are color-coded by red. (Figure 8)
- Directionality of the shape changes: Displayed by means of vector maps that show the shape differences between the mean shape of the two groups tested. (Figure 9)

- Significance maps: Showing the areas in where the shape changes were considered to be statistically significant. The p-values are color-coded showing blue where the raw p-value exceeds 0.05, that is to say the change for that point is not significant regardless the magnitude of it, and red in those points where the changes are significant. (Figure 10)

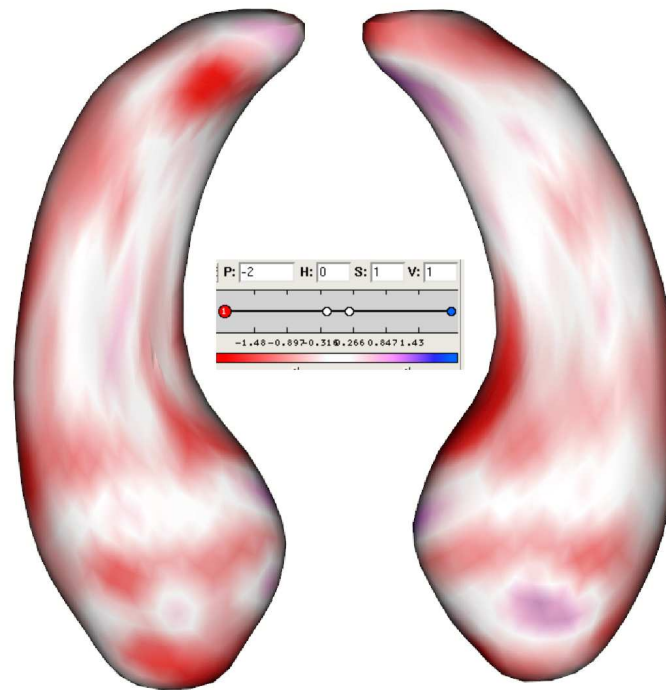


Figure 8: Magnitude and directional changes in hippocampus during the progression of schizophrenia.

5.3 Cross-sectional with interaction variables groupwise analysis scenario

This study was performed to determine the condylar morphological variation of osteoarthritic (OA) and asymptomatic temporomandibular joints (TMJ) and to determine its correlation with clinical signs and symptoms presented in the progression of the OA disease. The spectrum of clinical and pathological presentation of TMJ OA range from structural and functional failure of the joint with disc displacement and degeneration, to subchondral bone alterations (erosions), bone overgrowth (osteophytes), loss of articular fibrocartilage, and synovitis. Many osteoarthritic TMJs are associated with clinical symptoms [10] [12]. However, some are completely asymptomatic. The factor(s) which relates the morphological changes and clinical signs and symptoms (namely pain, joint sounds, and mandibular range of motion) is unknown.

Preliminary findings in this field were obtained using *shapeAnalysisMANCOVA*, computing Pearson correlation coefficients and their p-values between morphological variation and clinical variables for a dataset of 29 female patients with TMJ OA (Research Diagnostic Criteria for Temporomandibular Disorders Group III) and 40 female asymptomatic subjects. In this analysis scenario only results for correlations pain onset will be displayed, and were obtained with the following sequence of arguments,

Example command: 'infile - -numGroupTypes 1 - -numIndependent 1 - -interactionTest - -simpleCorrs'

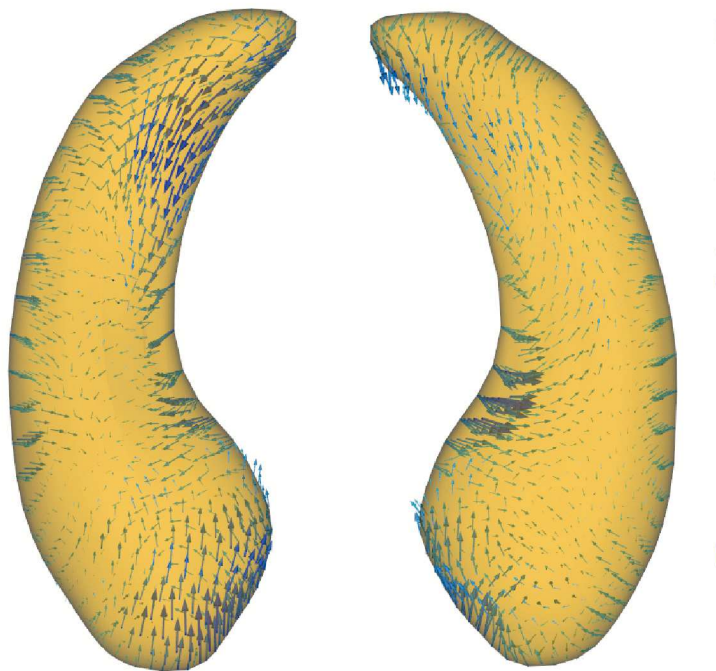


Figure 9: Directionality of shape changes displayed with vector maps.

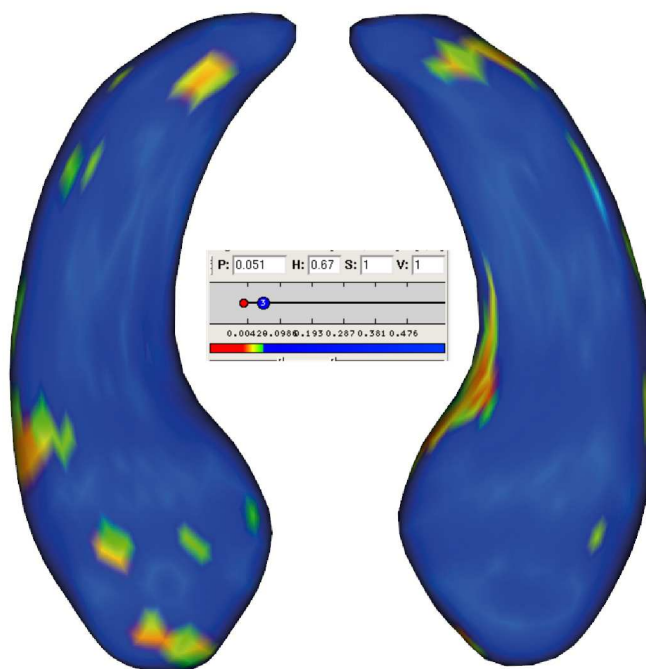


Figure 10: Statistical significance for the shape changes in hippocampus during the progression of schizophrenia.

Results

In the Pearson correlation coefficient maps, positive correlations are color-coded red and negative correlations are green. Yellow is used where there are no relevant correlation between shape and pain onset (coefficients equals or close to zero) P-values maps show the associated statistical significance for the correlation coefficients previously described. In the p-values maps very high correlations ($p < 0.01$) are color-coded with white, significant correlations are color-coded with red and green ($0.01 > p > 0.05$) and no significant correlations are color-coded with blue. The results for correlations between shape changes and pain onset are shown in figure 11.

6 Conclusions

We presented a novel statistical shape analysis tool called *shapeAnalysisMANCOVA*. The versatility of the tool allows it to be applied in many different shape morphometry applications. This paper described the methodology associated with the tool as well as a set of case studies in shape analysis. We are currently applying this tool in our studies of shape in schizophrenia, autism, TMJ and normal development.

Acknowledgements

This work was supported by grants from the NIDCR (DE018962), NAMIC (U54 EB005149) and NDRC (UNC P30 HD03110). Many thanks to the departments of Neurology and Psychiatry from Columbia University, and specially to Scott Schobel for providing the data presented in the Longitudinal groupwise analysis scenario (section 5.2).

References

- [1] C. Brechbuhler, G. Gerig, and O. Kubler. Parameterization of closed surfaces for 3d shape description. *Computer Vision, Graphics, Image Processing: Image Understanding*, 1995. 1, 2
- [2] M.K. Chung. *Statistical Morphometry in Neuroanatomy*. PhD thesis, McGill University, Montreal, 2001. 2.2
- [3] J.G. Csernansky, S. Joshi, L.E. Wang, J. Haller, M. Gado, J.P. Miller, U. Granander, and M.I. Miller. Hippocampal morphometry in schizophrenia via high dimensional brain mapping. *Proceedings of the National Academy of Sciences*, September 1998. 1
- [4] J.G. Csernansky, S. D.J. Wang, L. Jones, D. Rastogi-Cru, G. Posener, J.A. Heydebrand, J.P. Miller, U. Granander, and M.I. Miller. Hippocampal deformities in schizophrenia characterized by high dimensional brain mapping. *American Journal of Psychiatry*, December 2002. 1
- [5] C. Davatzikos, M. Vaillant, S. Resnick, J.L. Prince, S. Letovsky, and R.N. Bryan. A computerized method for morphological analysis of the corpus callosum. *Journal of Computer Assisted Tomography*, January 1996. 1
- [6] E. S. Edgington, editor. *Randomization Tests*. Academic Press, 1995. 2.1

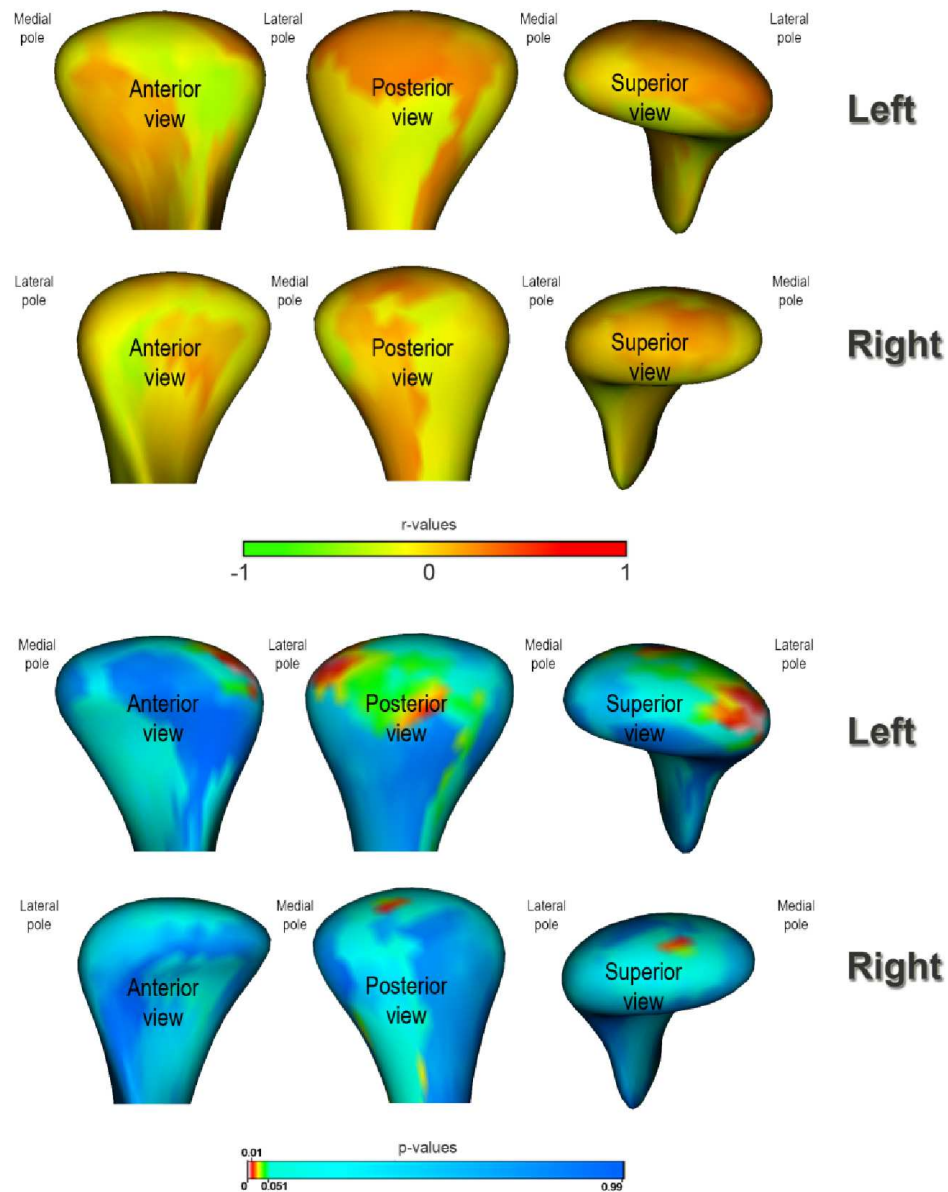


Figure 11: Morphological correlation with pain variables (pain onset). Up, Pearson correlation coefficients. Down, associated statistical significance maps. The significance maps show statistically significant correlations in the the lateral and posterior surfaces of the condyles.

- [7] C.R. Genovese, N.A. Lazar, and T. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 1(15):870–878, 2002. [2.2](#)
- [8] C.R. Genovese, N.A. Lazar, and T. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4):870–8, April 2002. [4.2](#)
- [9] P. Golland, W.E.L. Grimson, and R. Kikinis. Statistical shape analysis using fixed topology skeletons: Corpus callosum study. *Information Processing in Medical Imaging*, 1999. [1](#)
- [10] D.C. Hatcher. Craniofacial imaging. *Journal of the California Dental Association*, 6(19):27–34, 1991. [5.3](#)
- [11] H.C. Hazlett, M. Poe, and G. Gerig. Magnetic resonance imaging and head circumference study of brain size in autism: birth through age 2 years. *Archives of General Psychiatry*, 12(62):1366–1376, 2005. [5.1](#)
- [12] L.M. Helenius, D. Hallikainen, I. Helenius, J.H. Meurman, M. Knnen, and M. Leirisalo-Repo. Clinical and radiographic findings of the temporomandibular joint in patients with various rheumatic diseases: A case control study. *Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics*, 4(99):455–463, 2005. [5.3](#)
- [13] S. Joshi, M. Miller, and U. Grenander. On the geometry and shape of brain sub-manifolds. *Pattern Recognition and Artificial Intelligence*, 1997. [1](#)
- [14] T.E. Nichols and A.P. Holmes. Nonparametric analysis of pet functional neuroimaging experiments: A primer with examples. *Hum Brain Mapp*, 1(15):1–25, Jan 2001. [2.2](#)
- [15] I. Oguz, G. Gerig, S. Barre, and M. Styner. Kwmeshvisu: A mesh visualization tool for shape analysis. *Insight Journal*, July 2006. [4.2](#)
- [16] Cmake Org. www.cmake.org, 2009. [4](#)
- [17] D. Pantazis, R.M. Leahy, T.E. Nichol, and M. Styner. Statistical surface-based morphometry using a non-parametric approach. In *Int. Symposium on Biomedical Imaging(ISBI)*, pages 1283–1286, April 2004. [2.2](#)
- [18] D. Pantazis, T. E. Nichols, S. Baillet, and R. M. Leahy. Spatiotemporal localization of significant activation in meg using permutation tests. In *Information Processing in Medical Imaging*, pages 512–523, 2003. [2.2](#)
- [19] S. Pizer, D. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney. Segmentation, registration and measurement of shape variation via image object shape. *IEEE Transactions in Medical Imaging*, October 1999. [1](#)
- [20] R. Schlittgen. *Kevin Kim, Neil Timm: Univariate and Multivariate General Linear Models Theory and Applications with SAS; Second Edition*. Statistical Papers, March 2009. [3.2](#)
- [21] G.A.F. Seber. *Multivariate Observations*. Wiley Series in Probability and Statistics, August 2004. [2.1](#), [4.1](#)
- [22] E.R. Sowell, P.M. Thompson, C.J. Holmes, T. L. Jernigan, and A. W. Toga. In vivo evidence for post-adolescent frontal and striatal maturation. *Nature Neuroscience*, Sept./ Oct. 1999. [2.2](#)

- [23] M. Styner, G. Gerig, J.A. Lieberman, D. Jones, and D. Weinberger. Statistical shape analysis of neuroanatomical structures based on medial models. *Medical Image Analysis*, 2003. [1](#)
- [24] M. Styner, J.A. Lieberman, and G. Gerig. Boundary and medial shape analysis of the hippocampus in schizophrenia. *MICCAI 2003*, 2003. [1](#)
- [25] M. Styner, J.A. Lieberman, R.K. McClure, D.R. Weingberger, D.W. Jones, and G. Gerig. Morphometric analysis of lateral ventricles in schizophrenia and healthy controls regarding genetic and disease-specific factors. *Proceedings of the National Academy of Science, PNAS*, 102(12):4872–4877, 2005. [2.1](#)
- [26] M. Styner, J.A. Lieberman, D. Pantazis, and G. Gerig. Boundary and medial shape analysis of the hippocampus in schizophrenia. *Medical Image Analysis*, 8(3):197–203, 2004. [2.1](#)
- [27] M. Styner, I. Oguz, S. Xu1, C. Brechbuhler, D. Pantazis, J. Levitt, M. Shenton, and G. Gerig. Framework for the statistical shape analysis of brain structures using spharm-pdm. *Insight Journal*, July 2006. ([document](#)), [2](#), [5.1](#)
- [28] D. Thompson. *On Growth and Form*. Cambridge University Press, 1942. [1](#)
- [29] K. J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4:58–73, 1996. [2.2](#)
- [30] Y. Hochberg Y. Benjamini. Controlling false discovery rate: A practical and powerful approach to multiple testing. *Stat. Soc. Ser. B*, 1(57):289–300, 1995. [2.2](#)