QUANTIFYING AND TESTING HIPPOCAMPUS SHAPE DIFFERENCES: SCHIZOPHRENICS VS. CONTROLS

Emily O. Kistner and Keith E. Muller

Data from Gerig et al.

Department of Biostatistics
University of North Carolina at Chapel Hill
ekistner@bios.unc.edu Keith_Muller@UNC.EDU

OVERVIEW

1 INTRODUCTION

1.1 Motivation

Scientific Question

Do schizophrenics and controls differ in hippocampus shape?

Statistical Questions

1) How should we measure shape and shape change?

2) How should we analyze the measures?

Statistical Problems

1) Dealing with correlated observations.

2) Dealing with *zillions*[1] of correlated observations - a bigger problem.

3) Avoiding false positives and false negatives

[1] A statistical term indicating
independent sampling units $\gg$ # observations/unit

1.2 Confirmatory vs. Exploratory

Foundation of science:

a priori prediction and replication.

Need small type I error rate (min. false positives) & small type II error rate (min. false negatives).

Confirmatory: all hypotheses specified and fixed *a priori*, before data collected.

Exploratory: anything else.

Report clearly *all* of what done.
Badly misleading (dishonest) to pick and choose, and not report process.

1.3 We Are Reporting Exploratory Analysis

1 Volume (of hippocampus) differences,
and other data,
were analyzed and published previously.

2 Other exploratory analyses of hippocampus shape for these data have been conducted by
Gerig and Styner and reported in some venues.

3 An extensive set of thorough exploratory analyses, for a range of shape measures, were conducted by us (EOK and KEM).
We are reporting the “best”, and hence biased.

“If you torture your data long enough, they will tell you whatever you want to hear.”

See Muller, Barton, and Benignus (1984; available for download at http://www.bios.unc.edu/~muller under "publications")

2 ANALYSES CONSIDERED

2.1 Types of Data

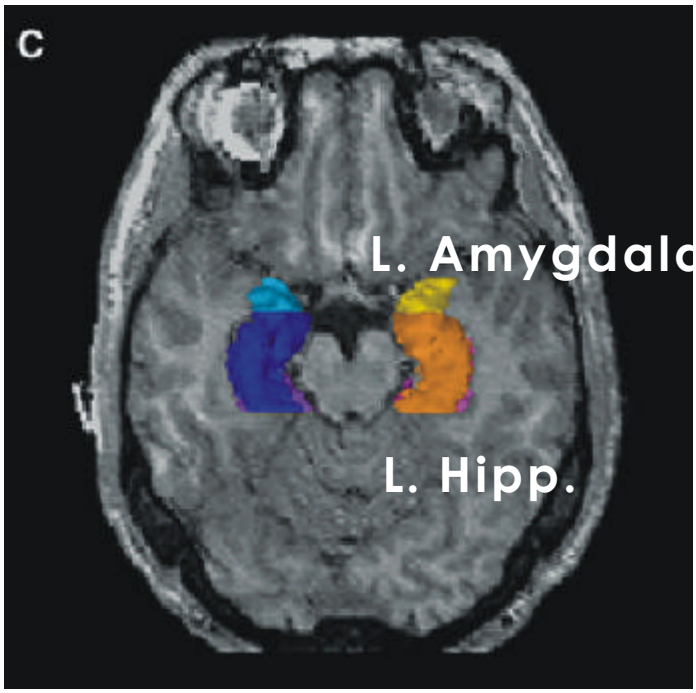


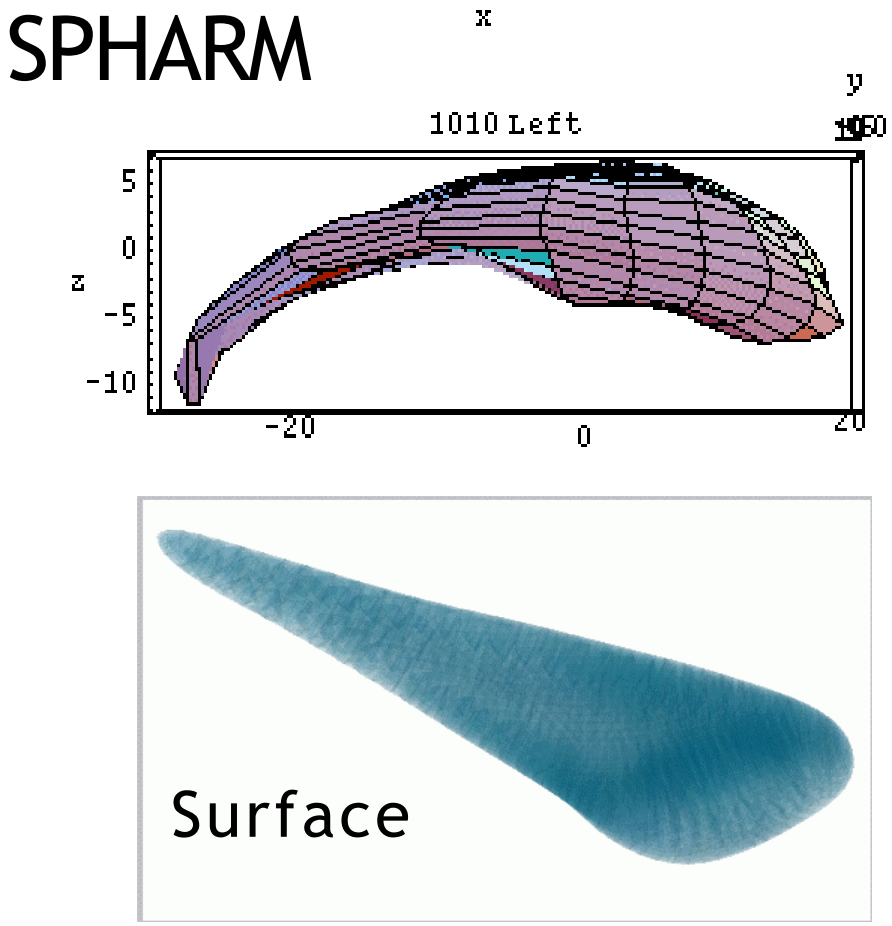
Figure 1 by Guido Gerig

A common origin and orientation were chosen, based on anatomic principles.

This is crucial in creating sensible data!

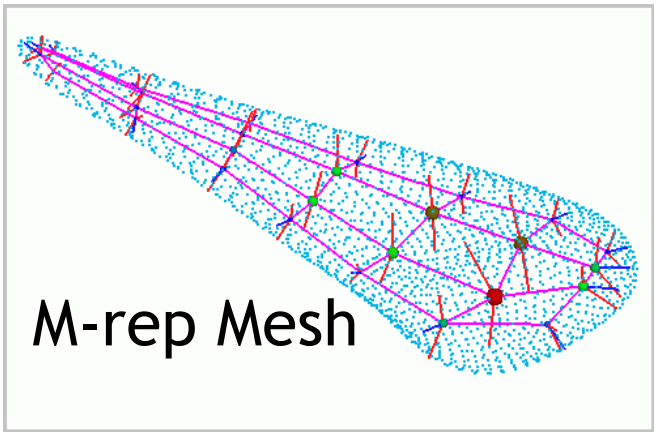
Unscaled data (did not adjust volume).
Analysis used here naturally separates scale.

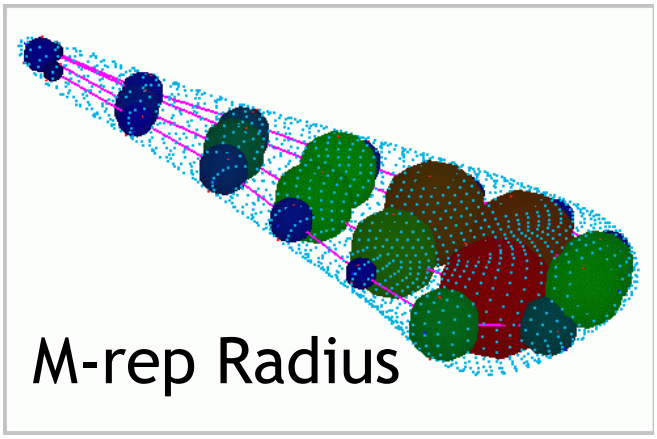
1 Spherical harmonics model generated smooth surface.
Data were x, y, and z of 12 "equally spaced" surface points per hippocampus side (Left, Right)



Figures 2.a and b by Guido Gerig

2 M-rep representation - x, y, z and radii of 24 points on an 8×3 grid for each hippocampus side.





Figures 3.a and b by Guido Gerig

2.2 Choice of Metric

Euclidean distance of surface point to origin: reduce x, y, and z to Distance.

Why?

+ Smaller number of variables in analysis easier to control false positive and false negative rate.

+ Summary measures usually more reliable (hence more sensitive).

+ Avoids need to model interrelationships among means of x, y, and z (and variances)

- May throw away information
if effect in only one dimension

2.3 Exploratory Analyses on Spherical Harmonics (Figure 2)

1a) Repeated Measures ANOVA
all trends (through 11th order polynomial) in point

1b) Spherical spatial covariance model
all trends (through 11th order polynomial) in point

We also worked with data based on 362 and 2252 surface points. We gradually realized not likely any additional information, given nature of the analysis.

The results are not reported in detail today.
Overall, only hints of small differences between control and Schizophrenics.

3 M-REP ANALYSIS

3.1 Data Reduction (Figure 3)

24 spheres per hippocampus per side; 48 per person

Two basic response variables:

Euclidean distance $= d(x, y, z)$ in mm from center of sphere to common origin

Radius in mm of each sphere

Each outcome analyzed completely separately.

Just one of many possible data reduction and analysis approaches.

3.2 Statistical Model

Data Structure

Multivariate Y (points within person correlated)

24 points in row, column order:
M11 M12 ... M18 M21 ... M28 M31 M32 ... M38
with column varying more quickly than row
(rows and columns from the M-rep grid).

Two such Y matrices (distance and radii), 79×48
79 subjects $\times$ 24 M-rep points $\cdot$ 2 sides (L/R)

Variables in Y (*within* subject effects) include

Side	of brain
Row	in M-rep structure
Column	in M-rep structure

Predictors in the model (*within* subject effects):
Side, Row, Column,
Side $\times$ Row, Side $\times$ Column, Row $\times$ Column,
Side $\times$ Row $\times$ Column (10+23+14 parameters)

Data Structure (continued)

Predictors in the model (*between* subject effects)

Age (years)

Drug Type (none, typical, or atypical)

Drug Type $\times$ Age (2 parameters) Interaction

Drug Type $\times$ Duration (months, 2 parameters)

Total *between* subject model parameters $= 8$

Model Reduction

Planned, fixed sequence of tests,
always from larger to smaller model,
with the order implied by the science.

Thanks to R. Hamer for very useful discussion.

There are $2^8 = 256$ possible between model.
We considered only 7 between models.
This greatly helps avoid false positive errors.

A similar statement holds for within model.

3.3 Results for Euclidean Distance

Residual analysis indicated no serious violation of the assumptions (≈ Gaussian homogeneous errors).

As an exploratory analysis, we examined both
a) Geisser-Greenhouse test ("Univariate approach to Repeated Measures," UNIREP) and
b)Wilks' test "Multivariate approach"

This may have introduced some optimistic bias.

We chose Geisser-Greenhouse.

Euclidean Distance (continued)

All tests involving Duration gave p-values > 0.05.
Hence Duration was dropped from between model.

All tests involving Side gave p-values > 0.05.
Hence Side was dropped from within model.

This is equivalent to averaging over Side, and reducing the Y matrix to a 79×24 matrix, with each column equal to the (left point + right point)$/2$.

After additional model reduction based on no differences between Typical and Atypical drug, the between subject model included
Age, Drug (yes, no), and Age $\times$ Drug (yes, no)

Test of Row $\times$ Col $\times$ Drug $\times$ Age
in the final model gave p = 0.0097 (GG test).

Row $\times$ Column shape was judged to be quadratic by quadratic (step-down interaction trend tests).

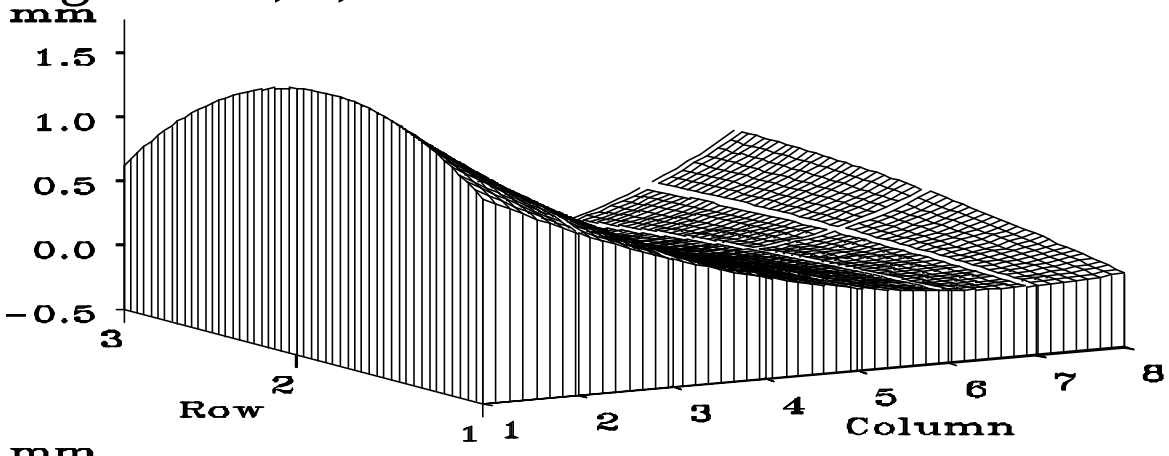
Euclidean Distance (continued)

Differences in hippocampus shape between patients and controls as measured by distance of M-rep spheres from origin are represented in Fig. 4a, b, c.

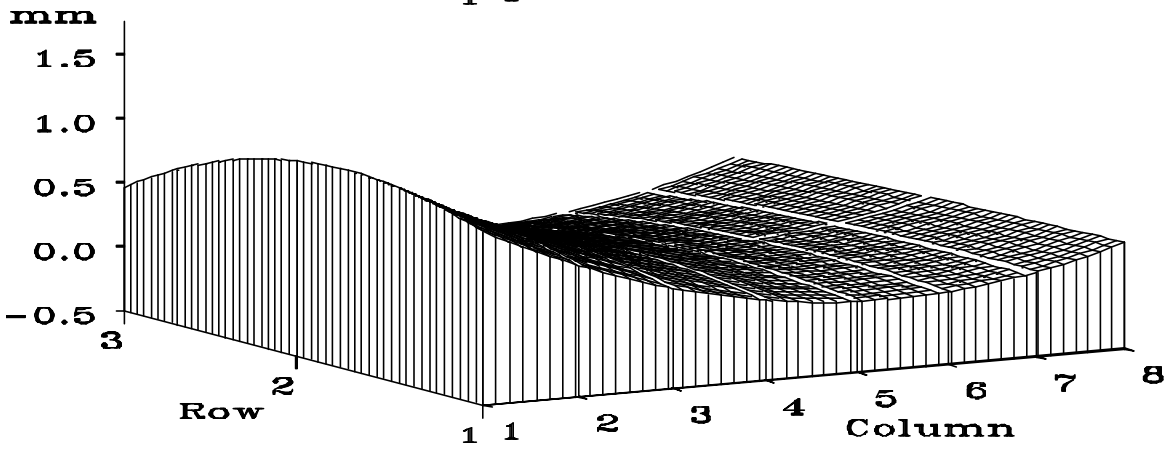
Quadratic (Row) $\times$ quadratic (Column) predicted surface.

The differences in distance between patients and controls *increases over time*, and are located mostly in the tail (and head).

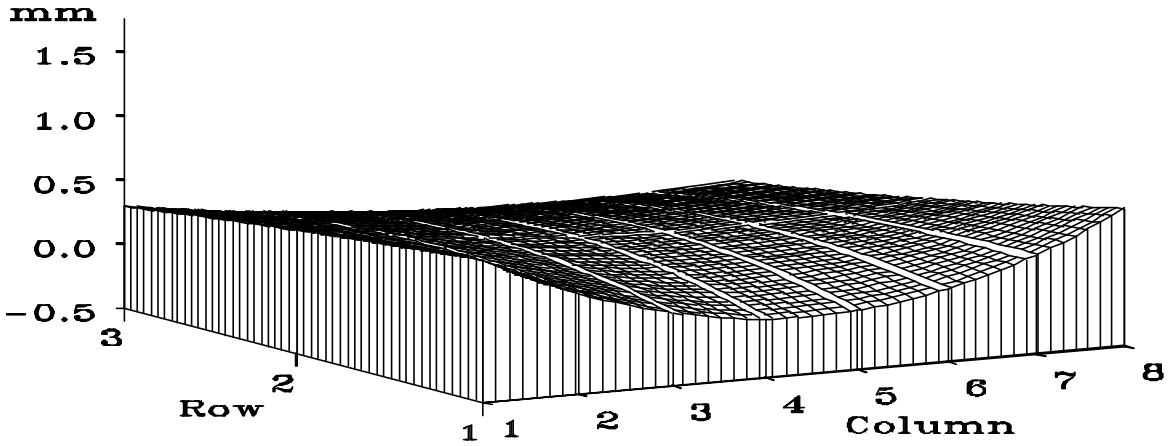
Figures 4a, b, c: Patient – Control Distance Diff.



4.a
Age
40



4.b
Age
30



4.c
Age
20

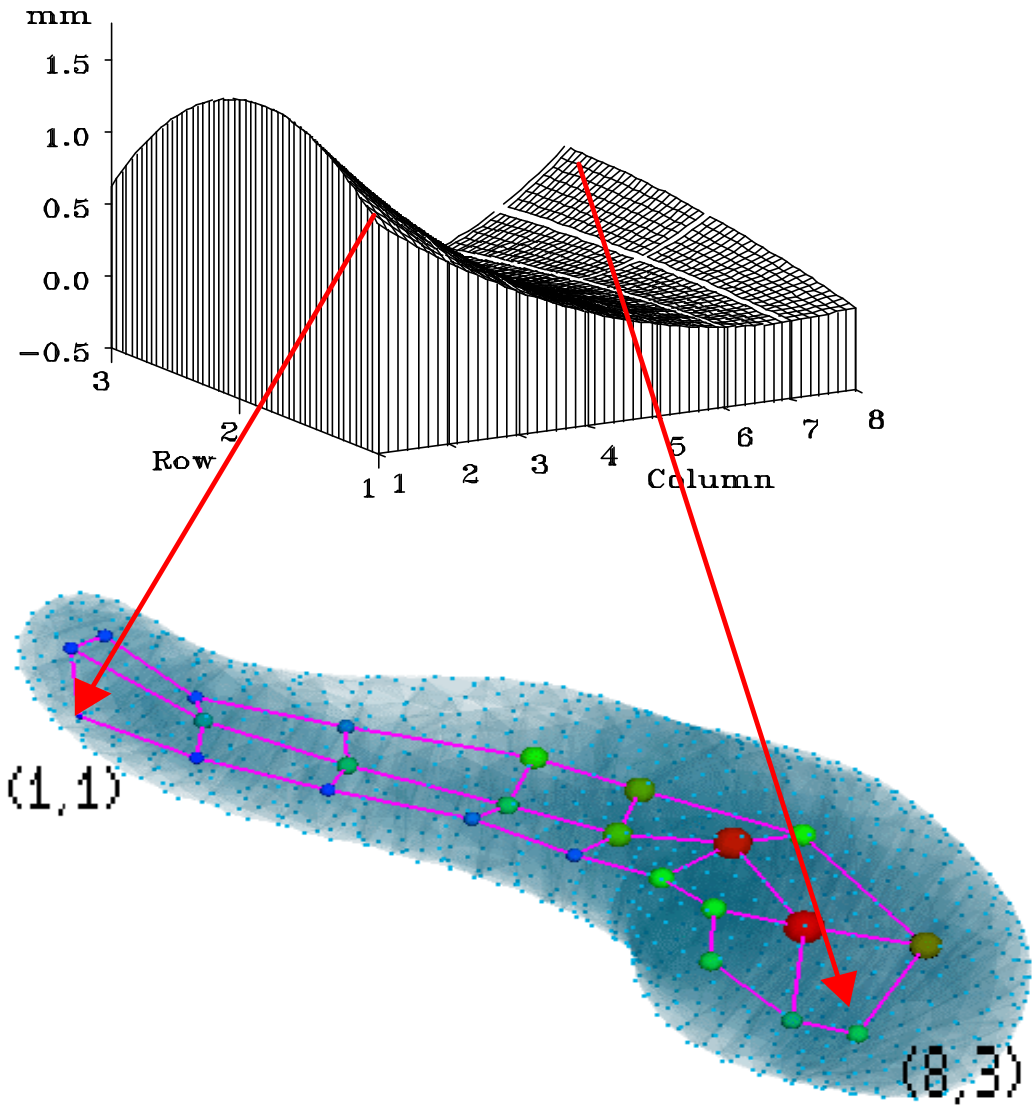


Figure 5 from Guido Gerig (modified)

Shape Change Due to Aging in Controls

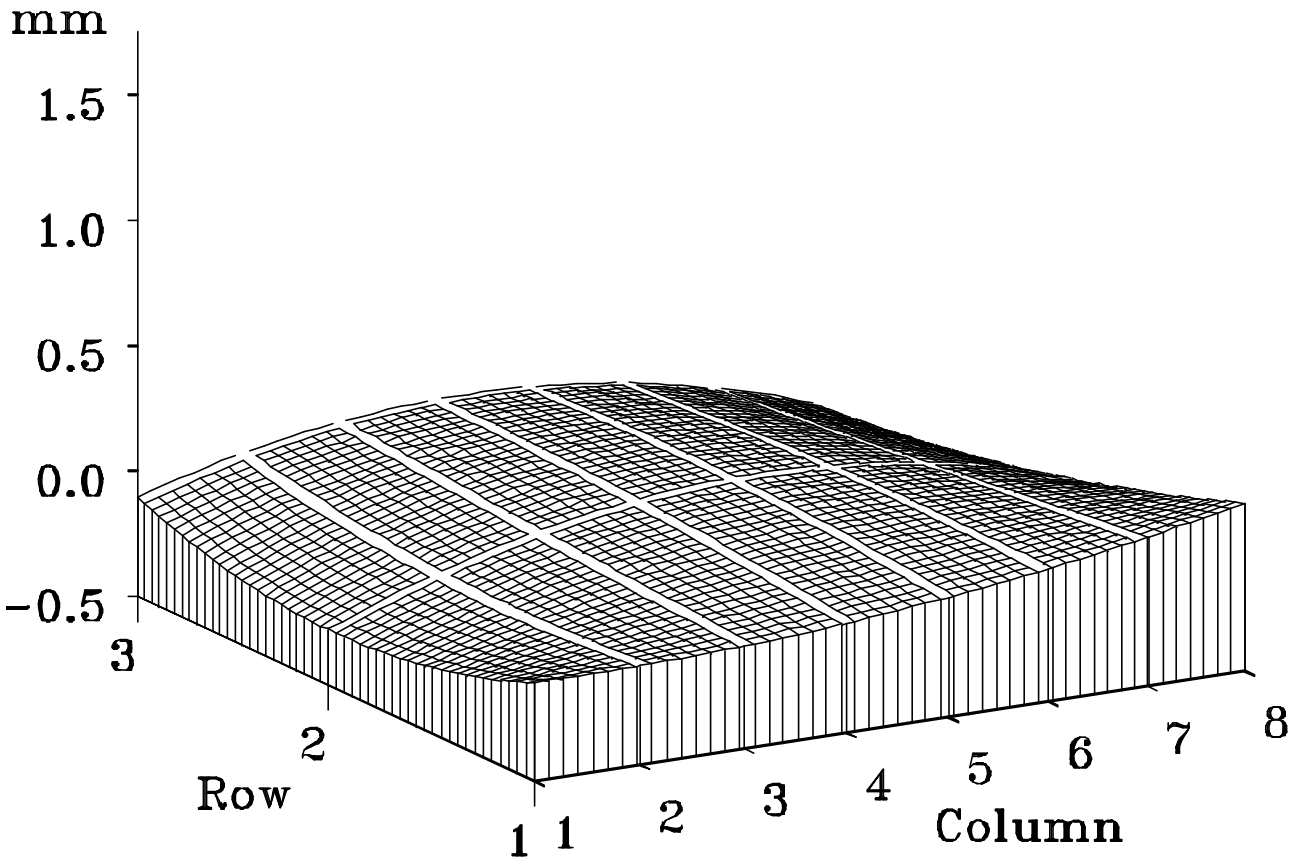


Figure 6: Distance Differences after 10 years, within Controls

3.4 Results for Radius

The residual analysis of radius raised some concern about violation of homogeneous Gaussian errors, including heavy tails and positive skewness.

Hence some transformations were considered. Logarithm (base 2, for convenience) of the radius made distributions less skewed and less heavy tailed.

Wilks' test was chosen rather than Geisser-Greenhouse test (due to value of diagnostic).

Radius (continued)

Model reduction gave final between model of
Age,
Drug Type (none, typical, atypical),
Age × Drug (yes, no), and
Duration × Drug (yes, no) as predictors.

Test of Side × Row × Col × Drug × Duration
gave p = 0.0077
Test of Side × Row × Col × Drug × Age
gave p = 0.0421

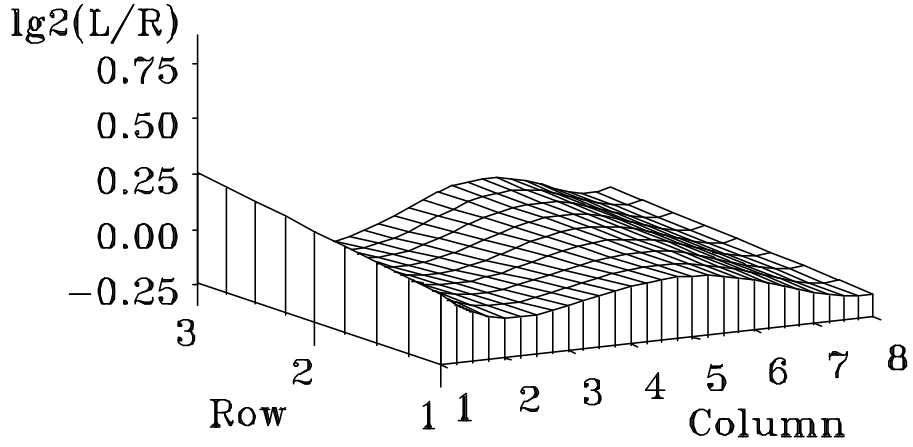
Row × Column shape judged to be linear by quintic (step-down interaction trend tests).

Radius (continued)

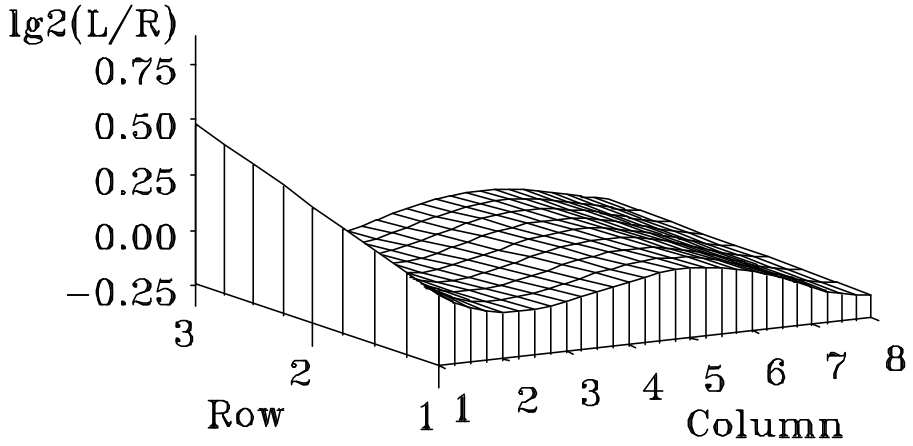
Differences in hippocampus shape *asymmetry*, measured by $\log_2(\text{M-rep radius Left/Right})$, are in Figures 7a,b,c for *Typical* drug group $-$ controls, Figures 7d,e,f for *Atypical* drug group $-$ controls.

Differences across (Row, Column) of a radius reflects shape information.
Points at the head and tail of the object, not the center, show differences.

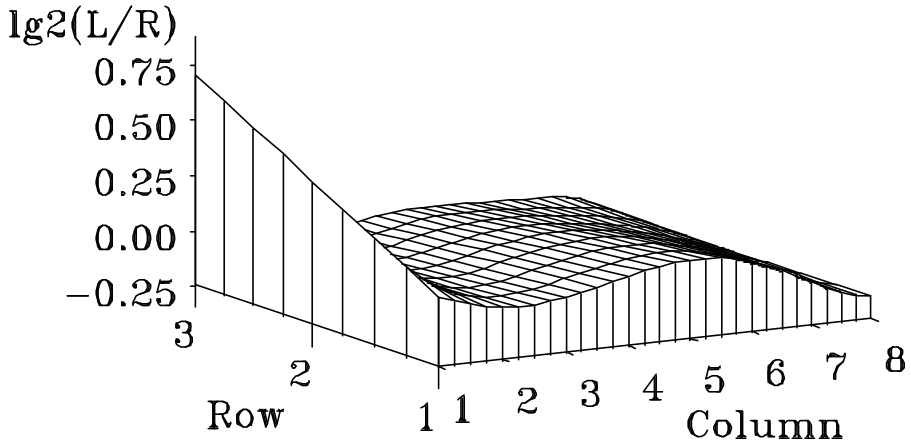
Fig. 7a,b,c Typcl $\log_2$(L/R)−Cntrl $\log_2$(L/R) radius



7.a
Age 40
Dur 16.7

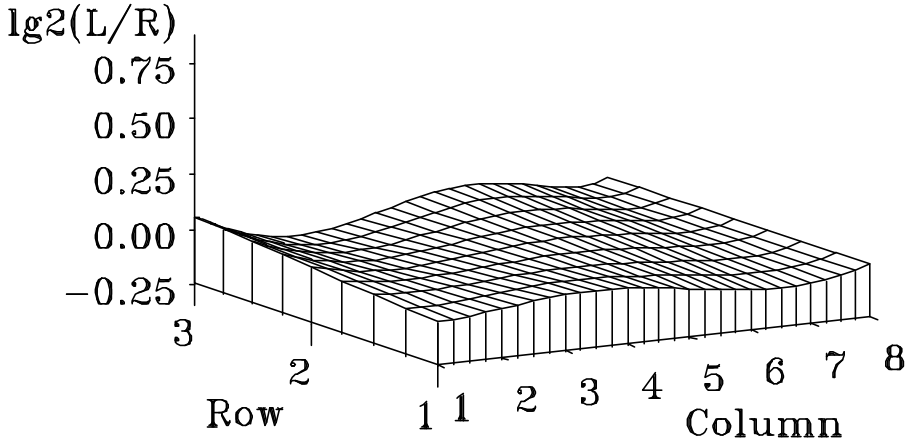


7.b
Age 30
Dur 8.3

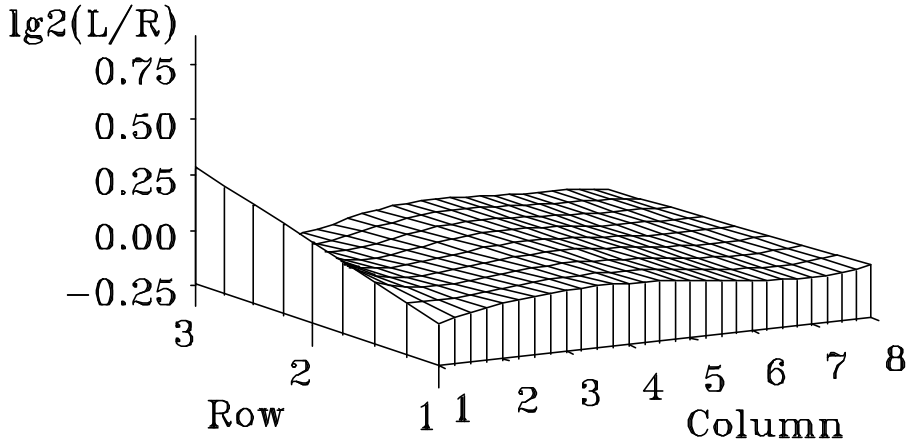


7.c
Age 20
Dur 0

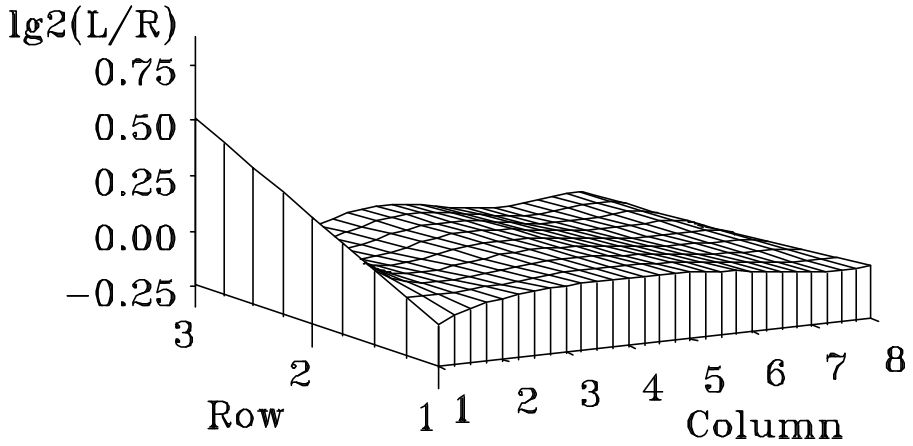
Fig. 7e,f,g Atyp $\log_2$(L/R)−Cntrl $\log_2$(L/R) radius



7.d
Age 40
Dur 16.7



7.e
Age 30
Dur 8.3



7.f
Age 20
Dur 0

The Atypical treated patients start (at an early age) less far from the Controls than do Typical treated. Is this a treatment effect or clinical selection bias?

Differences between patients and controls in hippocampus radius asymmetry
decrease over time.

Radius Asymmetry Change With Age in Cntrl

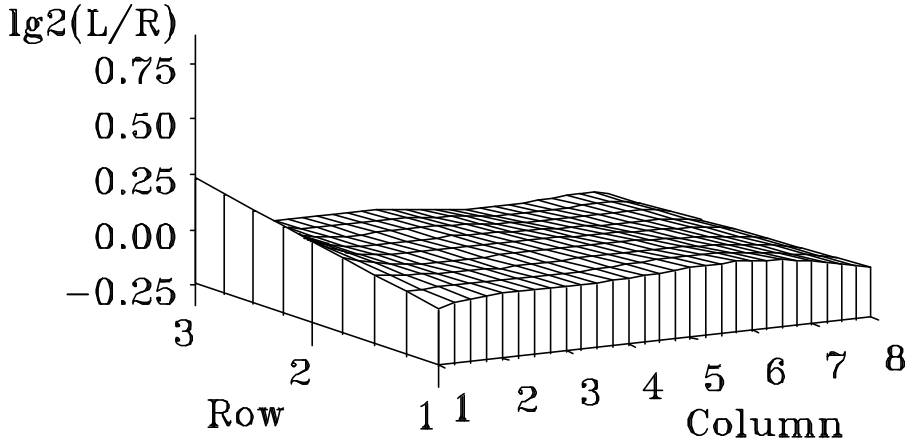


Figure 8: 10 year change in asymmetry in Controls

$$2^{.25} \approx 1.19 \Leftrightarrow$$

19% increase (?) in asymmetry per decade

4 DISCUSSION

4.1 Cautions.
We need replication.
Must report exploratory context of results.

4.2 Our choice of metrics in M-rep is
one choice of many possible.

4.3 Additional exploratory analysis using mixed models techniques to assess spatial covariance structure has appeal.

4.4 Why did M-rep work well?
Method focuses on image “central tendencies.”
We hypothesize that such “heavy” features intrinsically have smaller error variability, and hence more sensitivity.

APPENDIX

The General Linear Multivariate Model used:

$$\underset{79 \times 48}{\boldsymbol{Y}} = \underset{(79 \times 8)(8 \times 48)}{\boldsymbol{XB}} + \underset{79 \times 48}{\boldsymbol{E}}$$

Rows of $\boldsymbol{Y}$ and $\boldsymbol{X}$, and $\boldsymbol{E}$ for 79 subjects

48 cols in $\boldsymbol{Y}$, $\boldsymbol{E}$ M-Rep location:Side · Row · Col

Left			Right		
Row1	Row2	Row3	Row1	Row2	Row3
C1…C8	C1…C8	C1…C8	C1…C8	C1…C8	C1… C8
$[\boldsymbol{y}_1$		$\boldsymbol{y}_{24}$	$\boldsymbol{y}_{25}$		$\boldsymbol{y}_{48}]$

$$\boldsymbol{Y} = [\, \boldsymbol{y}_1 \quad \dots \quad \boldsymbol{y}_{48} \,]$$

$$\underset{79 \times 8}{\boldsymbol{X}} = \begin{bmatrix} \mathbf{1}_{n_{\mathrm{C}}} & \underline{\mathrm{Age}}_{\mathrm{C}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{1}_{n_{\mathrm{T}}} & \underline{\mathrm{Age}}_{\mathrm{T}} & \mathbf{1}_{n_{\mathrm{T}}} & \mathbf{0} & \underline{\mathrm{Age}}_{\mathrm{T}} & \mathbf{0} & \underline{\mathrm{Dur}}_{\mathrm{A}} & \mathbf{0} \\ \mathbf{1}_{n_{\mathrm{A}}} & \underline{\mathrm{Age}}_{\mathrm{A}} & \mathbf{0} & \mathbf{1}_{n_{\mathrm{A}}} & \mathbf{0} & \underline{\mathrm{Age}}_{\mathrm{A}} & \mathbf{0} & \underline{\mathrm{Dur}}_{\mathrm{T}} \end{bmatrix}$$

Int. Age Typ Atyp Drg×Age Drg×Dur

$$\underset{8 \times 24}{\boldsymbol{B}} = \begin{bmatrix} \beta_{0\ \mathrm{L,Row1,Col1}} & \cdots\cdots & \beta_{0\ \mathrm{R,Row3,Col8}} \\ \beta_{\mathrm{Age_C}\ \mathrm{L,Row1,Col1}} & \cdots\cdots & \beta_{\mathrm{Age_C}\ \mathrm{R,Row3,Col8}} \\ \vdots & & \vdots \\ & & \\ & & \\ & & \\ & & \\ \beta_{\mathrm{D}\times\mathrm{D_T}\ \mathrm{L,Row1,Col1}} & \cdots\cdots & \beta_{\mathrm{D}\times\mathrm{D_T}\ \mathrm{R,Row3,Col8}} \end{bmatrix}$$

The General Linear Hypothesis is H_0: $\boldsymbol{CBU} = \mathbf{0}$
$\boldsymbol{C}$ gives between subject contrasts (rows of $\boldsymbol{B}$)
$\boldsymbol{U}$ gives within subject contrasts (columns of $\boldsymbol{B}$)

Sources Tested

Age
Drug
Drug × Age
Drug × Duration

Side
Row
Column
Side × Row
Side × Column
Row × Side
Side × Row × Column

All between crossed with all within, including
Drug × Age × Side × Row × Column
Drug × Duration × Side × Row × Column