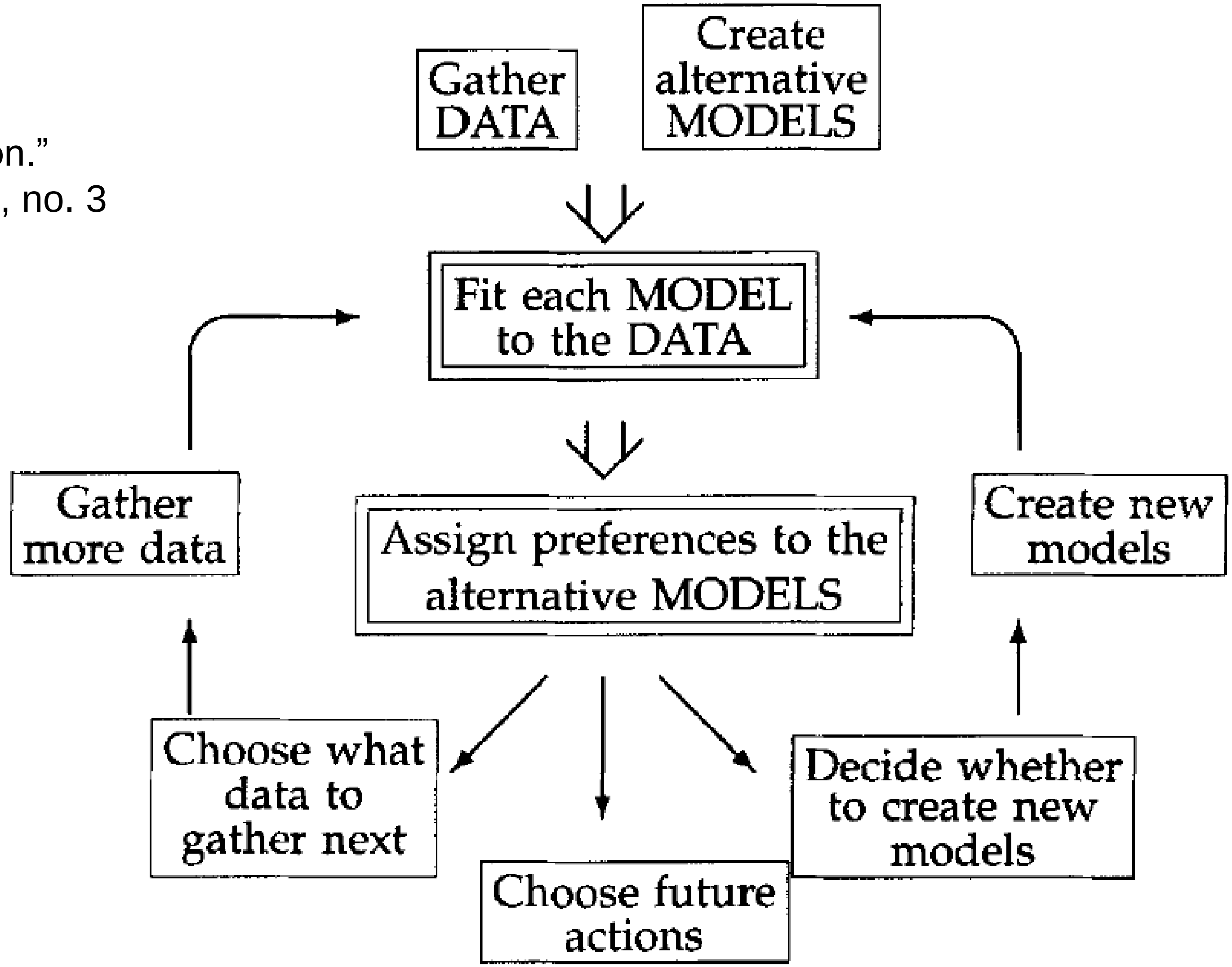
Data mining for neuroimaging data

John Ashburner

MODELLING

The Scientific Process

MacKay, David JC. "Bayesian interpolation." Neural computation 4, no. 3 (1992): 415-447.



Model Selection

- Search for the best of a number of models:

  $p(\mathbf{Y}|M_0)$, $p(\mathbf{Y}|M_1)$, $p(\mathbf{Y}|M_2)$, $p(\mathbf{Y}|M_3)$, $p(\mathbf{Y}|M_4)$…

- Cross-validation is essentially hypothesis testing.
  - Learn a hypothesis/model from the training data.
  - Test it on the data that was left out.

- Other model selection strategies are also possible – eg Bayesian Model Selection.

- The complexity of the best model depends on how much data is available.

Models of brain data

- Currently in neuroimaging, most models are for single voxels (ie mass-univariate).
  - Lots of separate models
  - Assumes independence among voxels.
  - Simple interpretation of differences

- The alternative is to model all the data.
  - Multivariate
  - More difficult to interpret

- Simplifying principles may emerge from more complex models.

UNIVARIATE OR MULTIVARIATE?

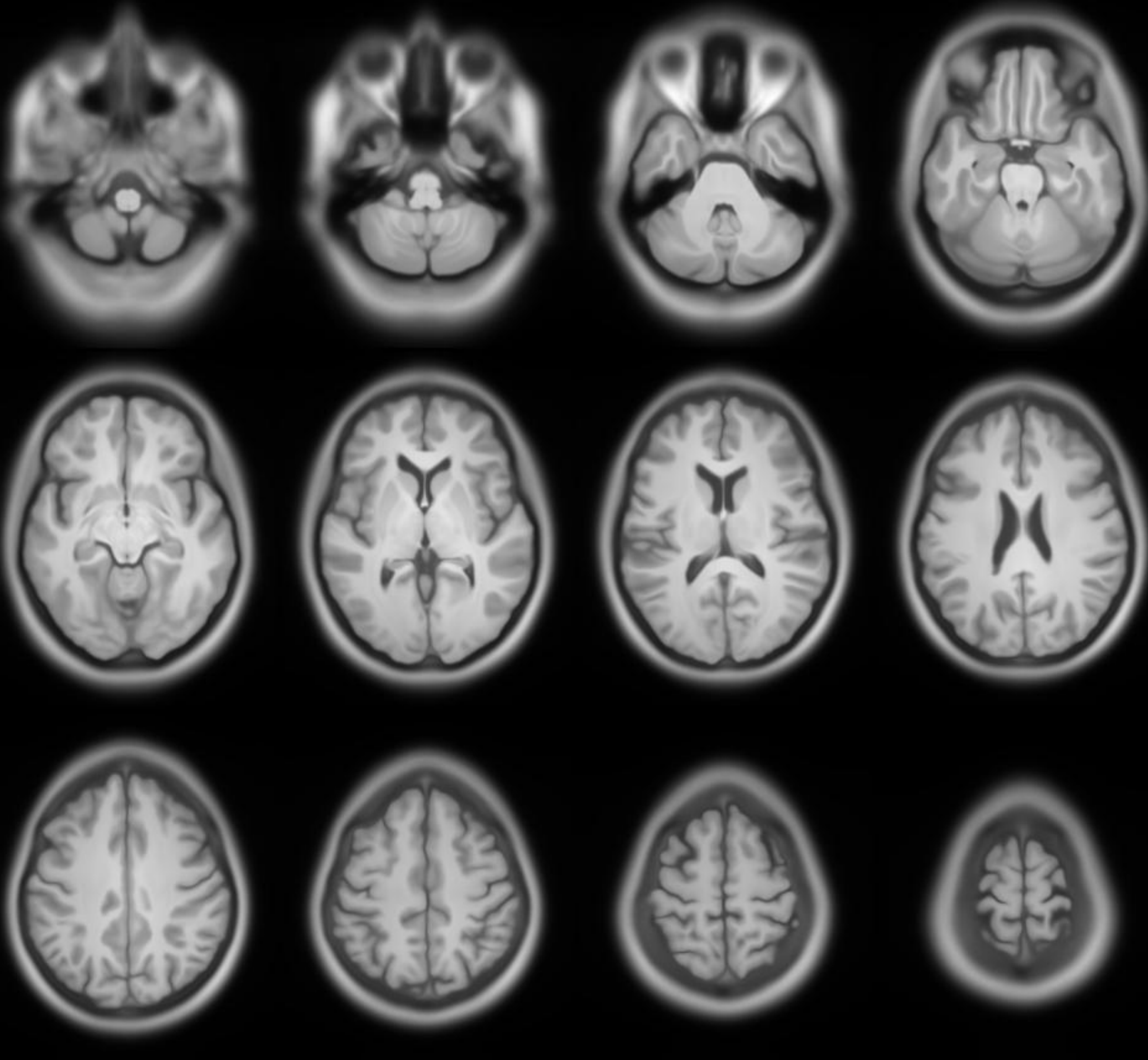
Are biological structures multivariate?

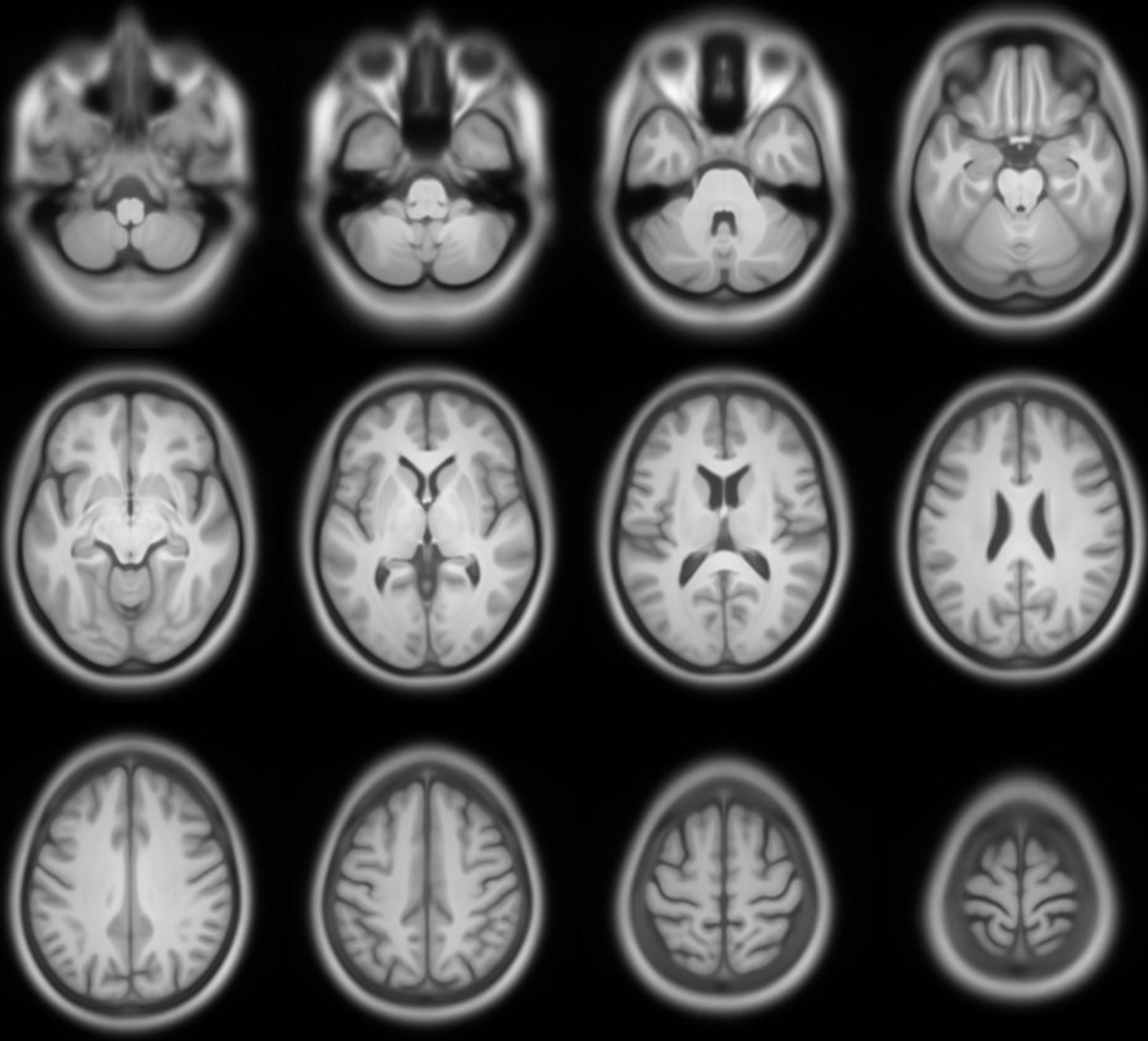
- Eventual shape is a result of growth.
- Growth is a result of gene expression and other factors.
- Each gene may be expressed in more than one voxel.

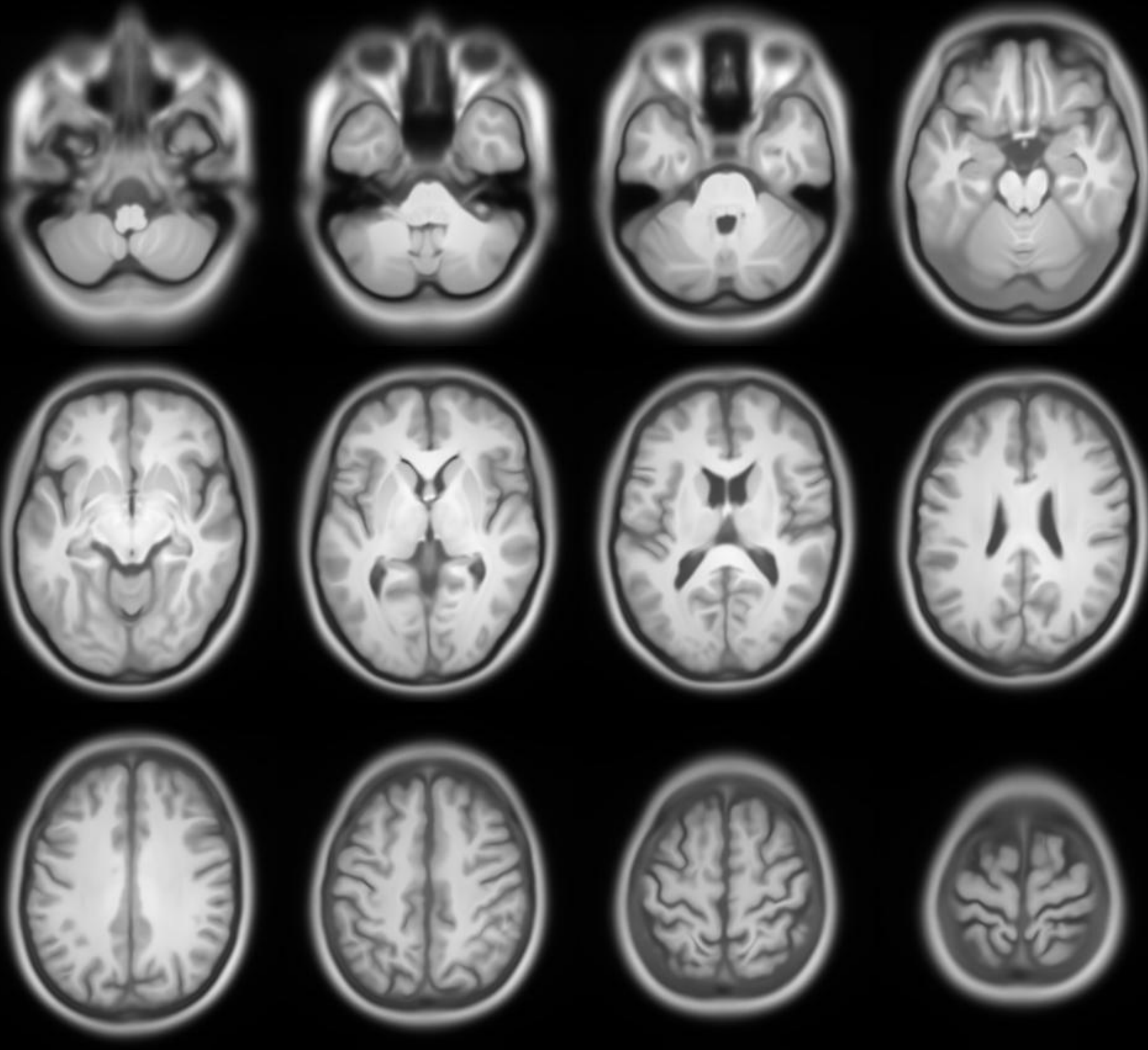
- We have known for a long time that, eg, left leg length is correlated with right leg length.

  It is, however, far more necessary to bear in mind that there are many unknown laws of **correlation of growth**, which, when one part of the organisation is modified through variation, and the modifications are accumulated by natural selection for the good of the being, will cause other modifications, often of the most unexpected nature (C. Darwin, 1859).

Can male-female differences be localised?

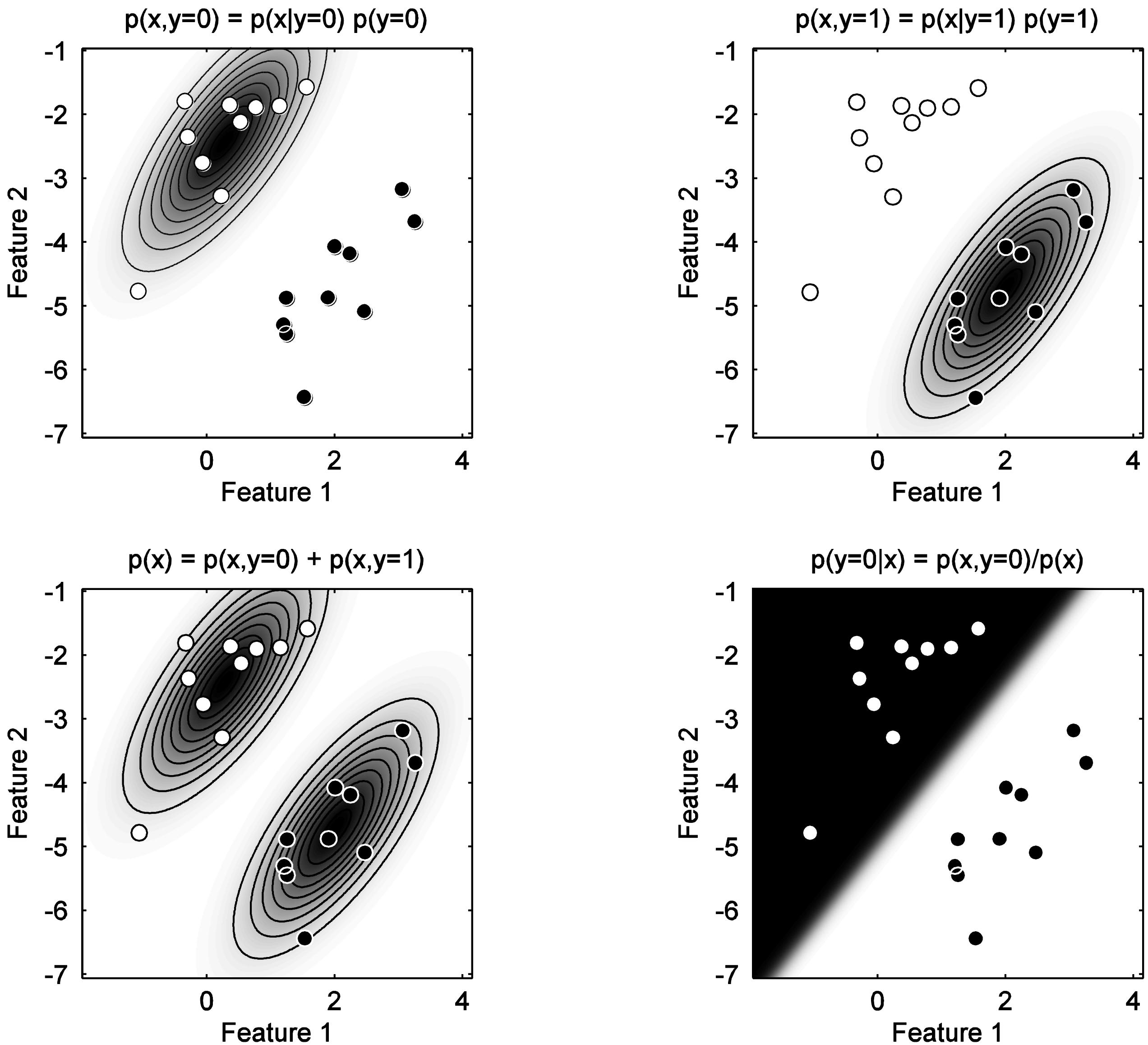






SOME MULTIVARIATE METHODS

Fisher's Linear Discriminant



Generative Model for Discrimination

- Generative:

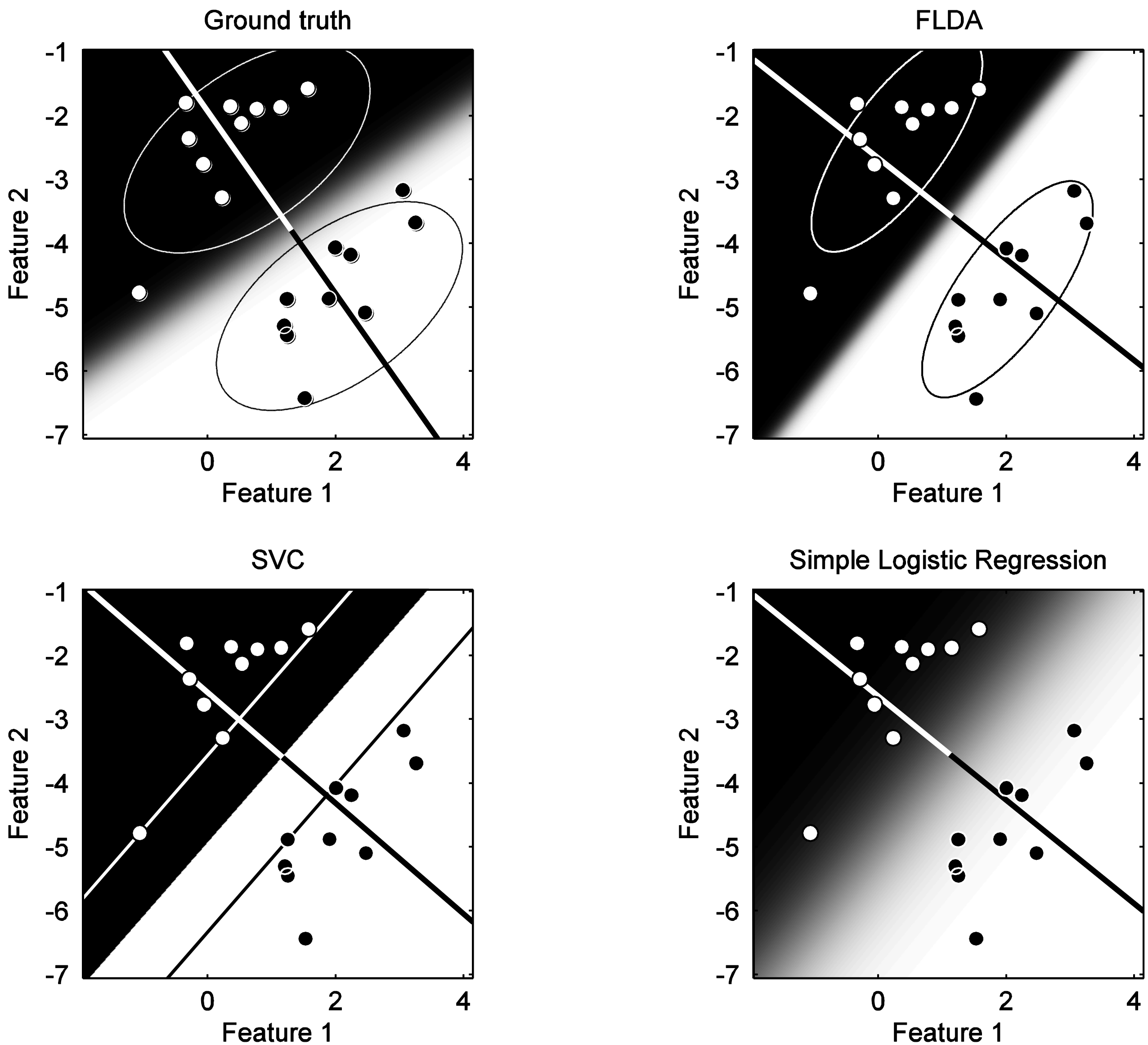
$$P(t=1|\mathbf{x}) = \frac{p(\mathbf{x}|t=1)P(t=1)}{p(\mathbf{x}|t=0)P(t=0) + p(\mathbf{x}|t=1)P(t=1)}$$

Where $\mathbf{x}$ feature data

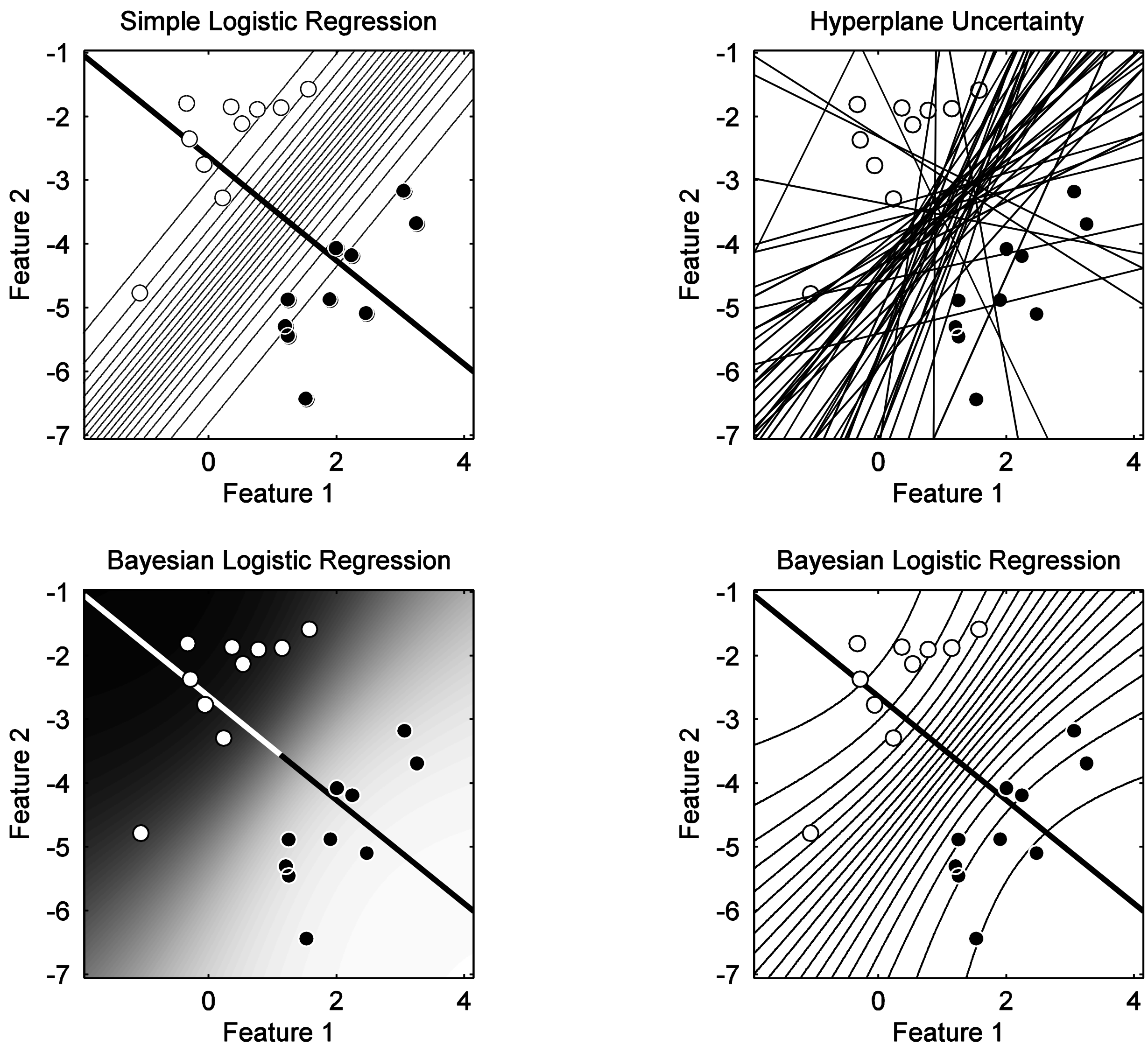
t prediction

- Discriminative:
  - Directly learns to give $P(t=1|\mathbf{x})$
  - We are not normally interested in all the variables needed to represent within-group variability.
  - Only after a discriminative direction.

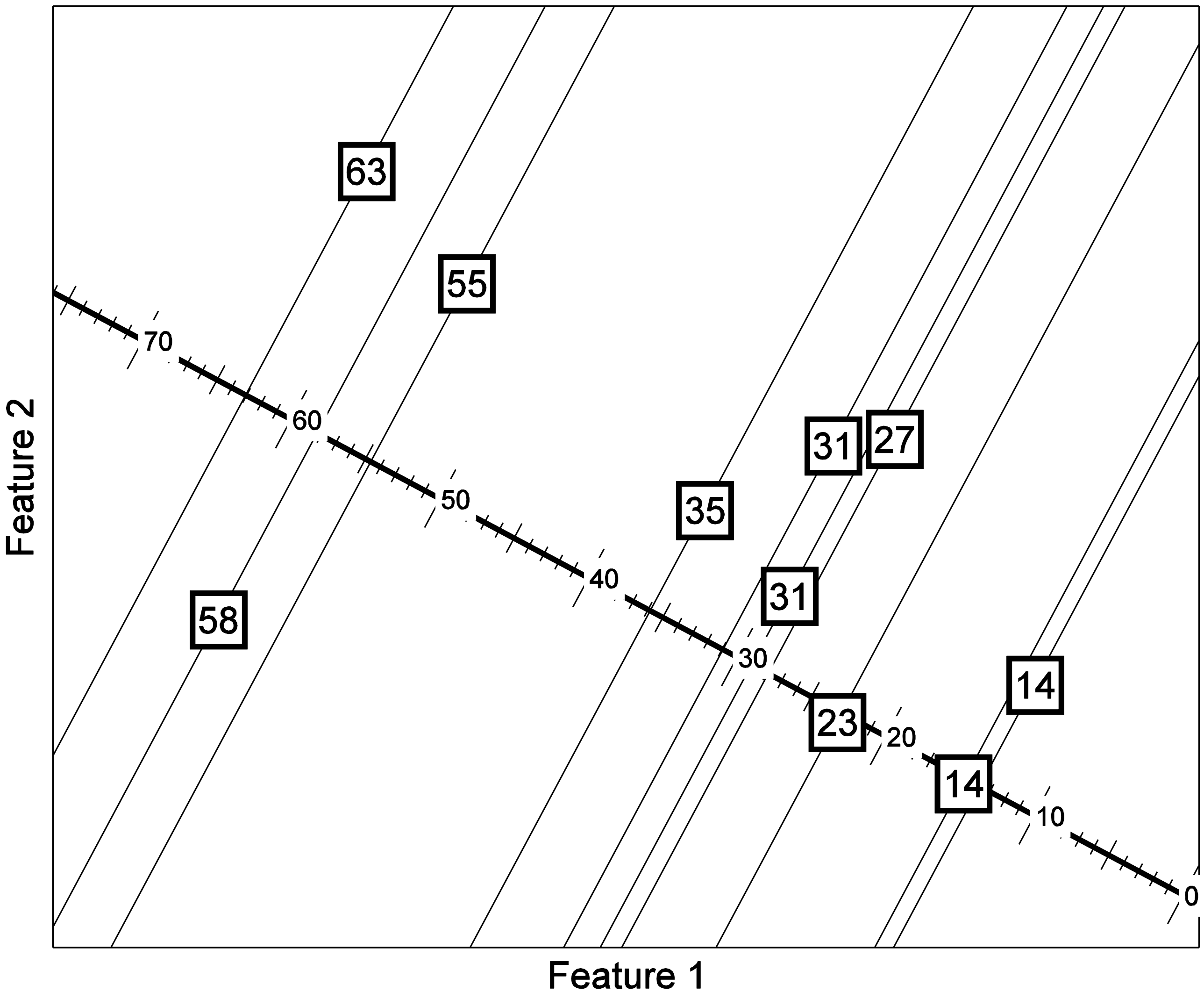
Linear Discrimination



Probabilistic Approaches



Regression



63
55
58
35
31
27
31
23
14
14
70
60
50
40
30
20
10
0
Feature 2
Feature 1

Gaussian Processes for Machine Learning

Book by Rasmussen & Williams available for free online at:

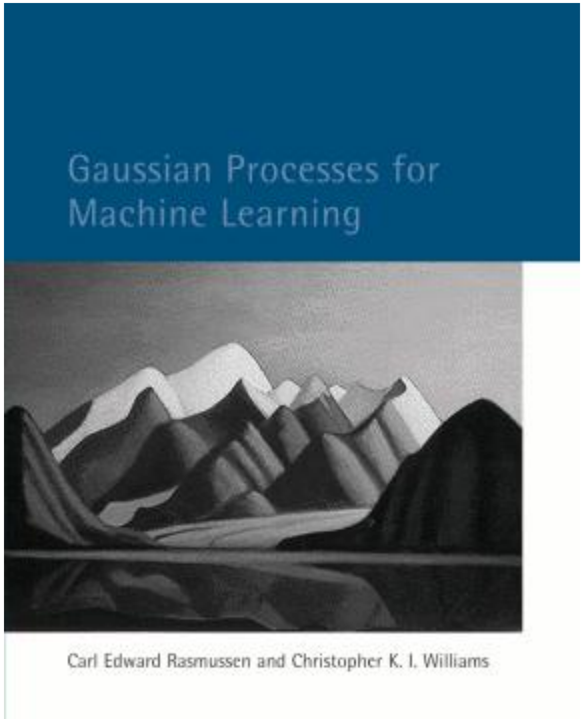
http://www.gaussianprocess.org/gpml

MATLAB code for regression and classification is also available.

Regression is relatively simple, but two approaches to classification are included.

- Laplace Approximation
- Expectation Propagation – more accurate

A Variational Bayes approach to GP classification is described in Bishop's PRML book.



TYPES OF FEATURES

Ugly Duckling Theorem

- An argument asserting that classification is impossible without some sort of bias.

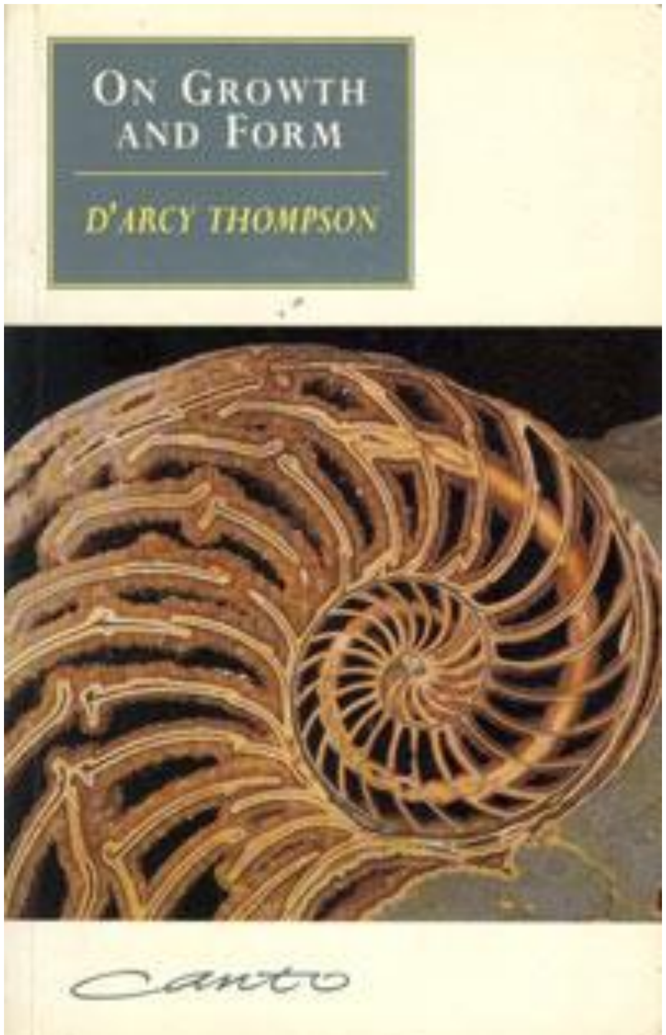
  Watanabe, Satosi (1969). Knowing and Guessing: A Quantitative Study of Inference and Information. New York: Wiley. pp. 376–377.

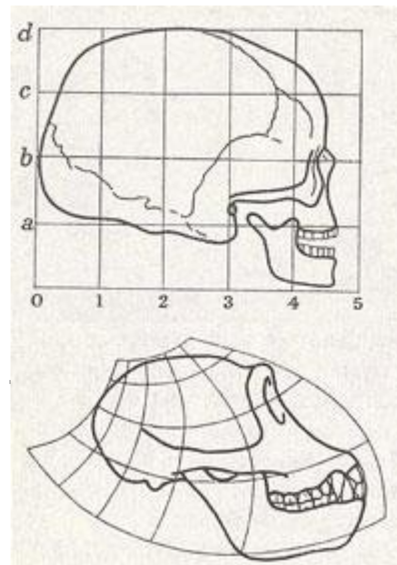
7.6. THEOREM OF THE UGLY DUCKLING

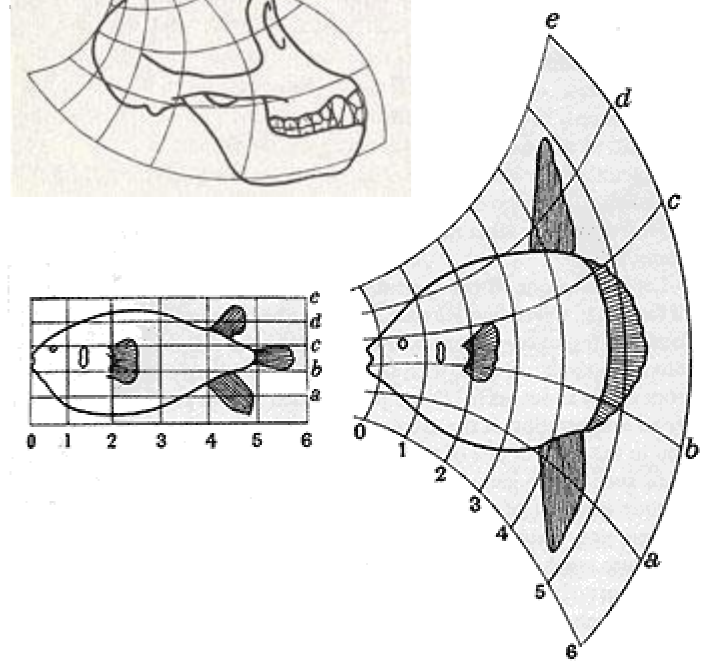
The purposes of this section is to show that from the formal point of view there exists no such thing as a class of similar objects in the world, insofar as all predicates (of the same dimension) have the same importance. Conversely, if we acknowledge the empirical existence of classes of similar objects, it means that we are attaching nonuniform importance to various predicates, and that this weighting has an extralogical origin.

How would the data be preprocessed to reveal useful features?

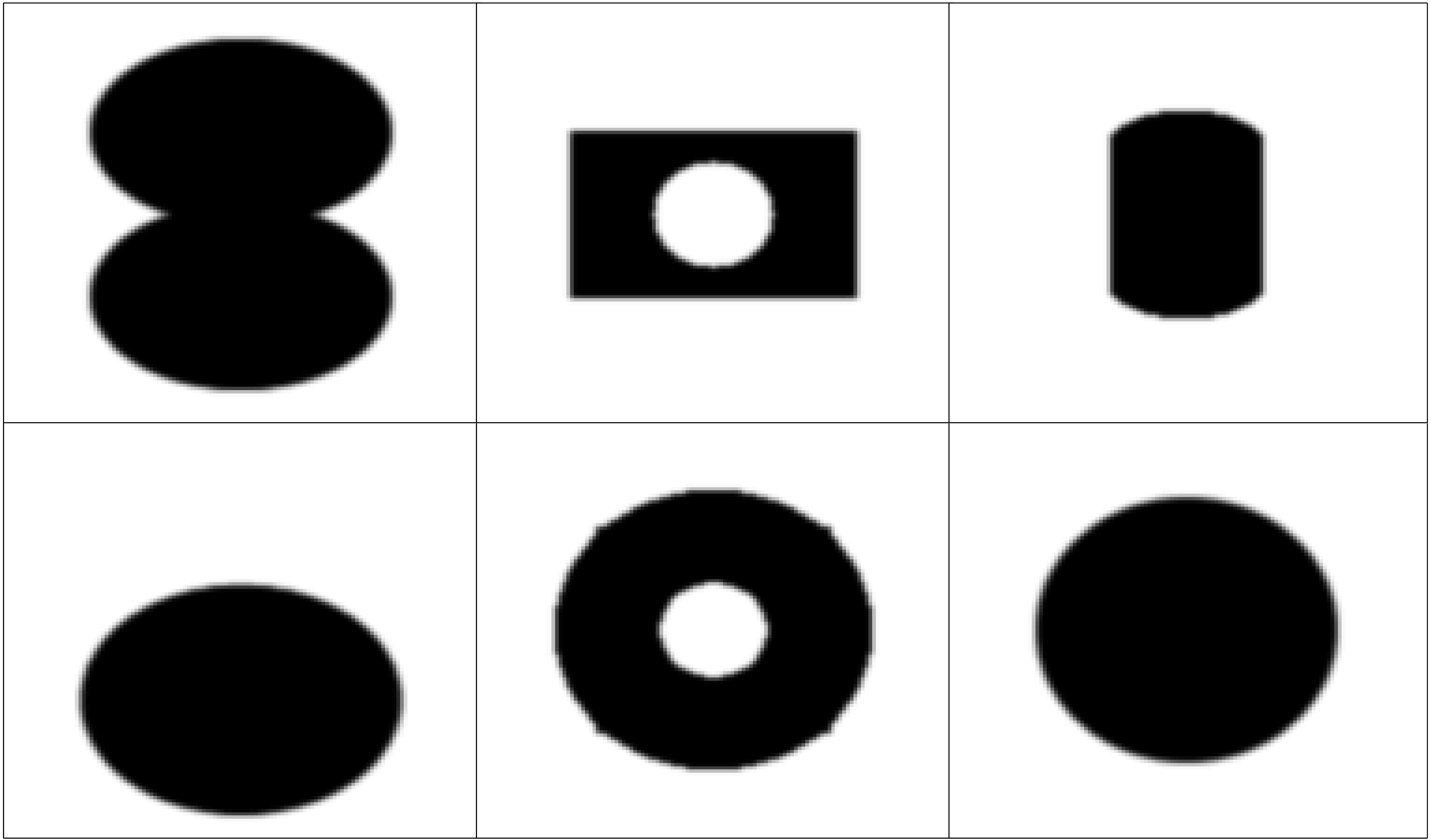
D’Arcy Thompson



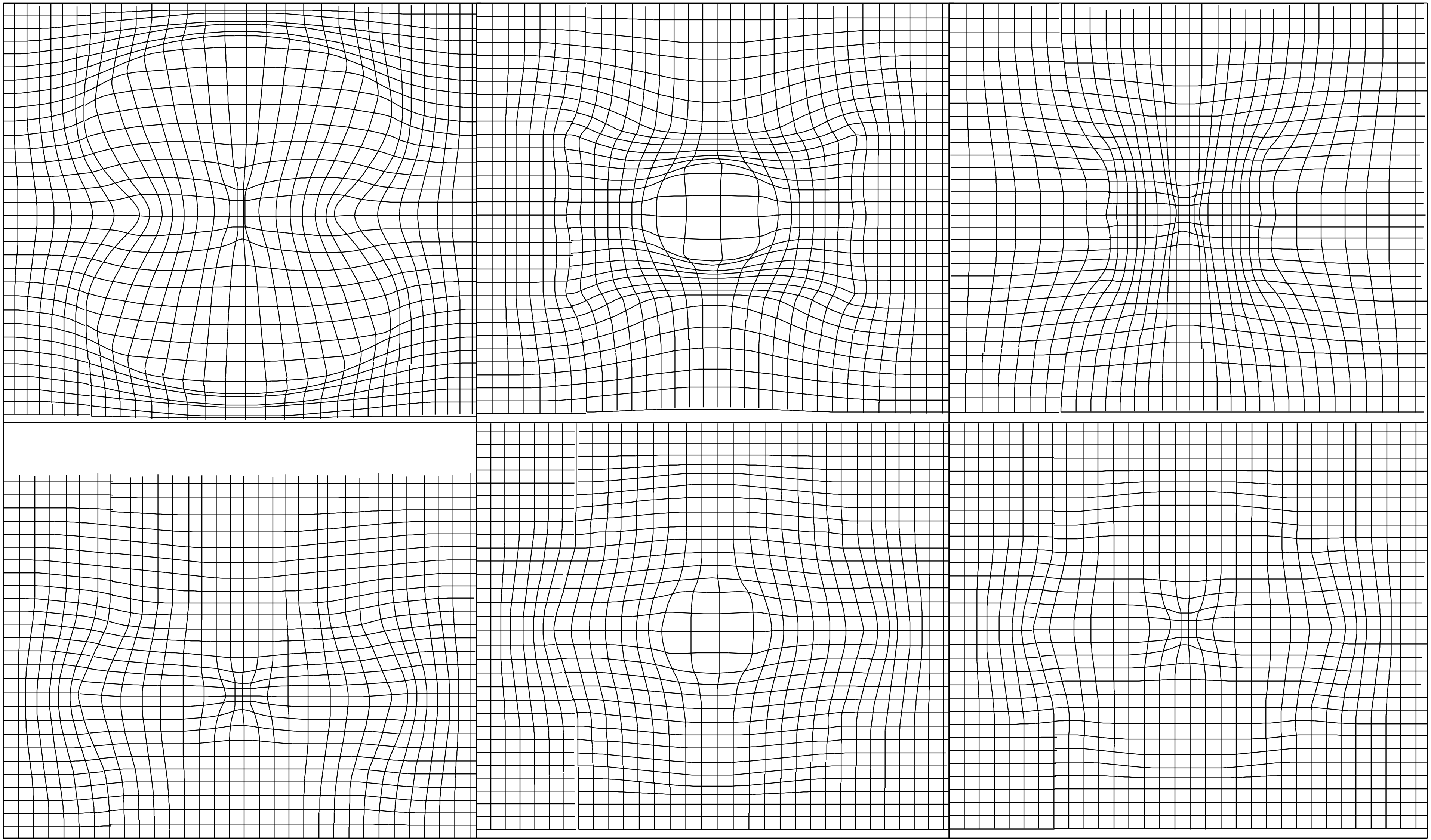




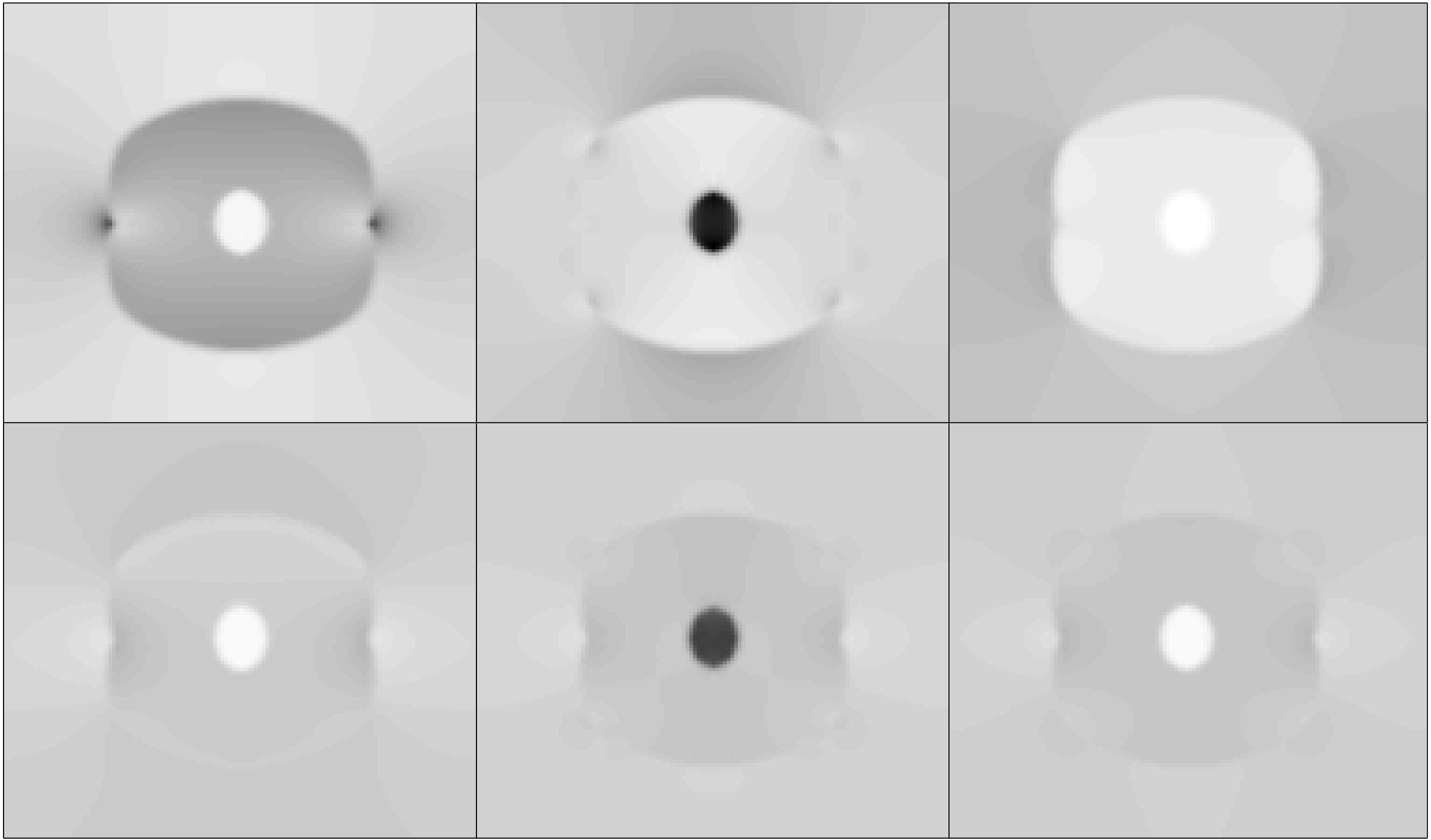
2D Example



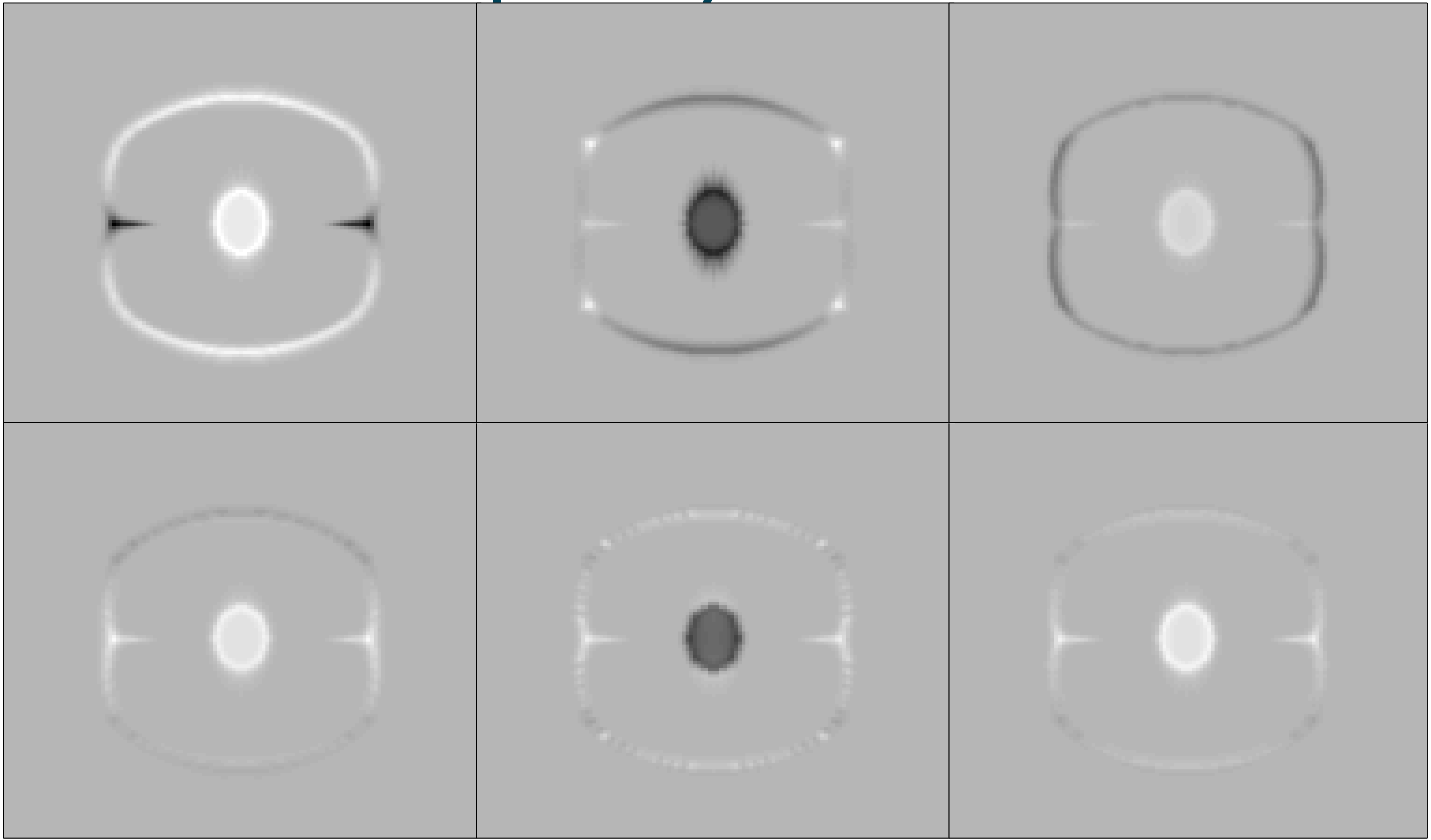
Diffeomorphic Deformations

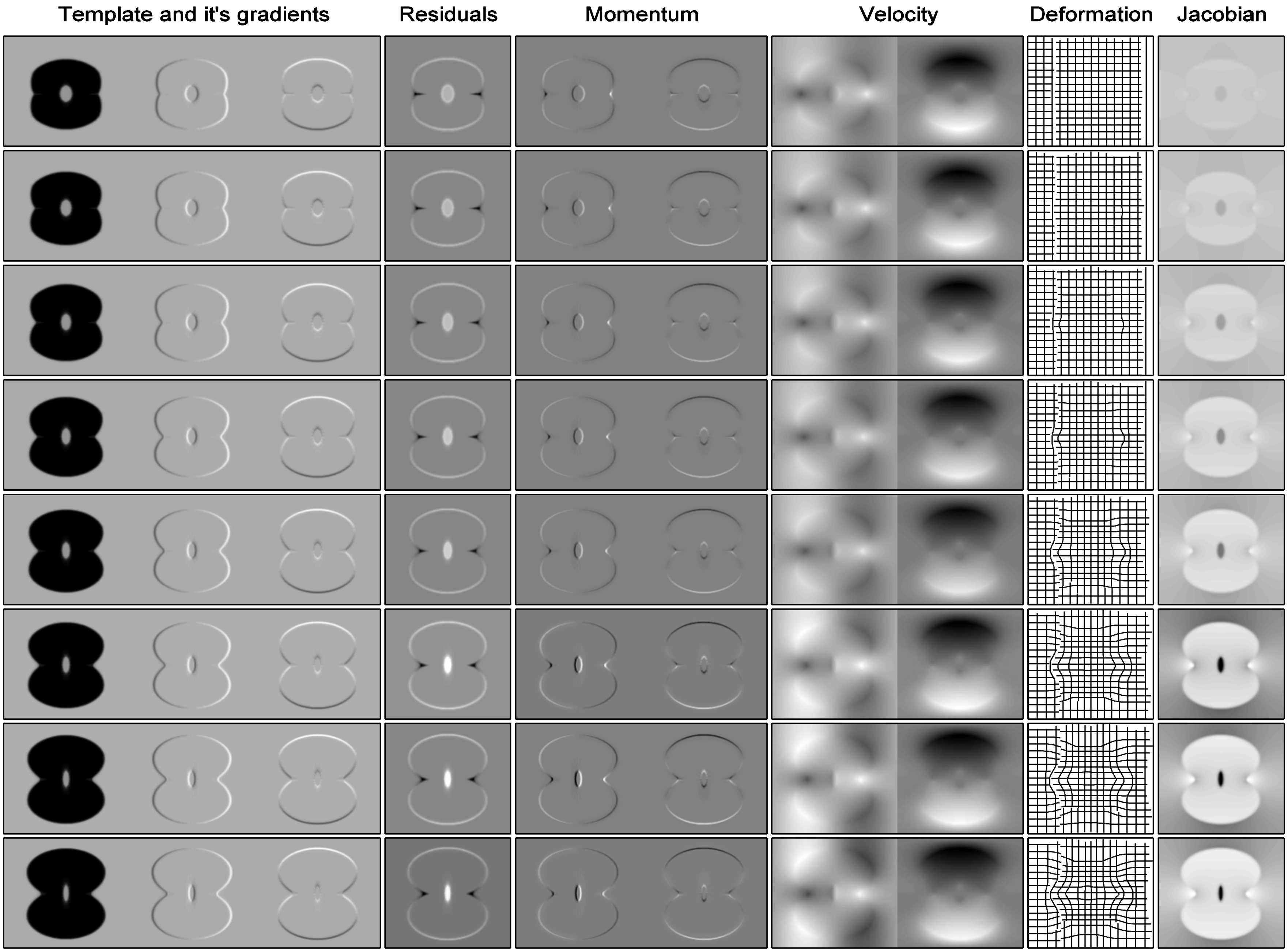


Jacobian Determinants (relative volumes)

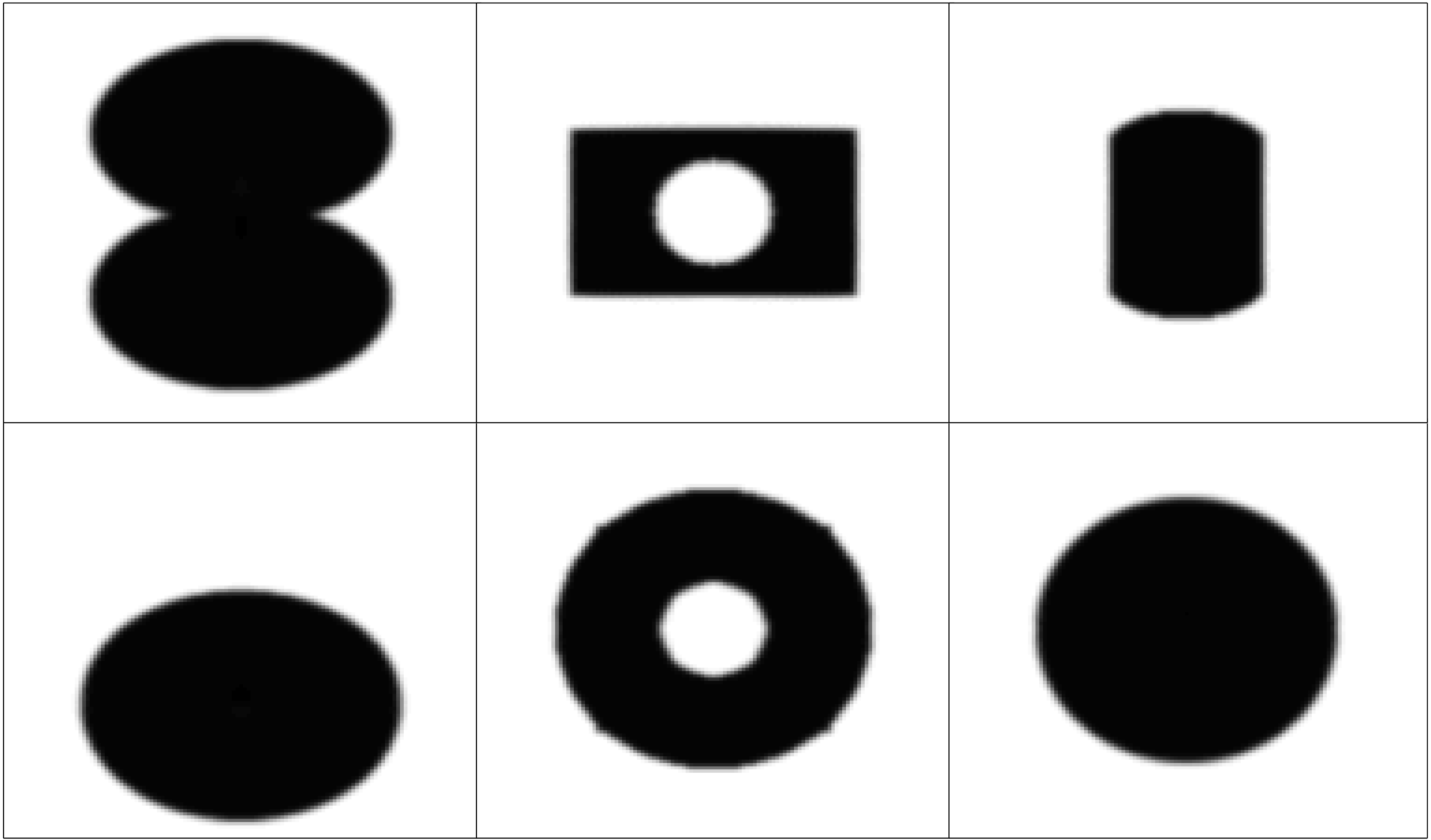


“Scalar Momenta”(Jacobian scaled residuals from diffeomorphisms)

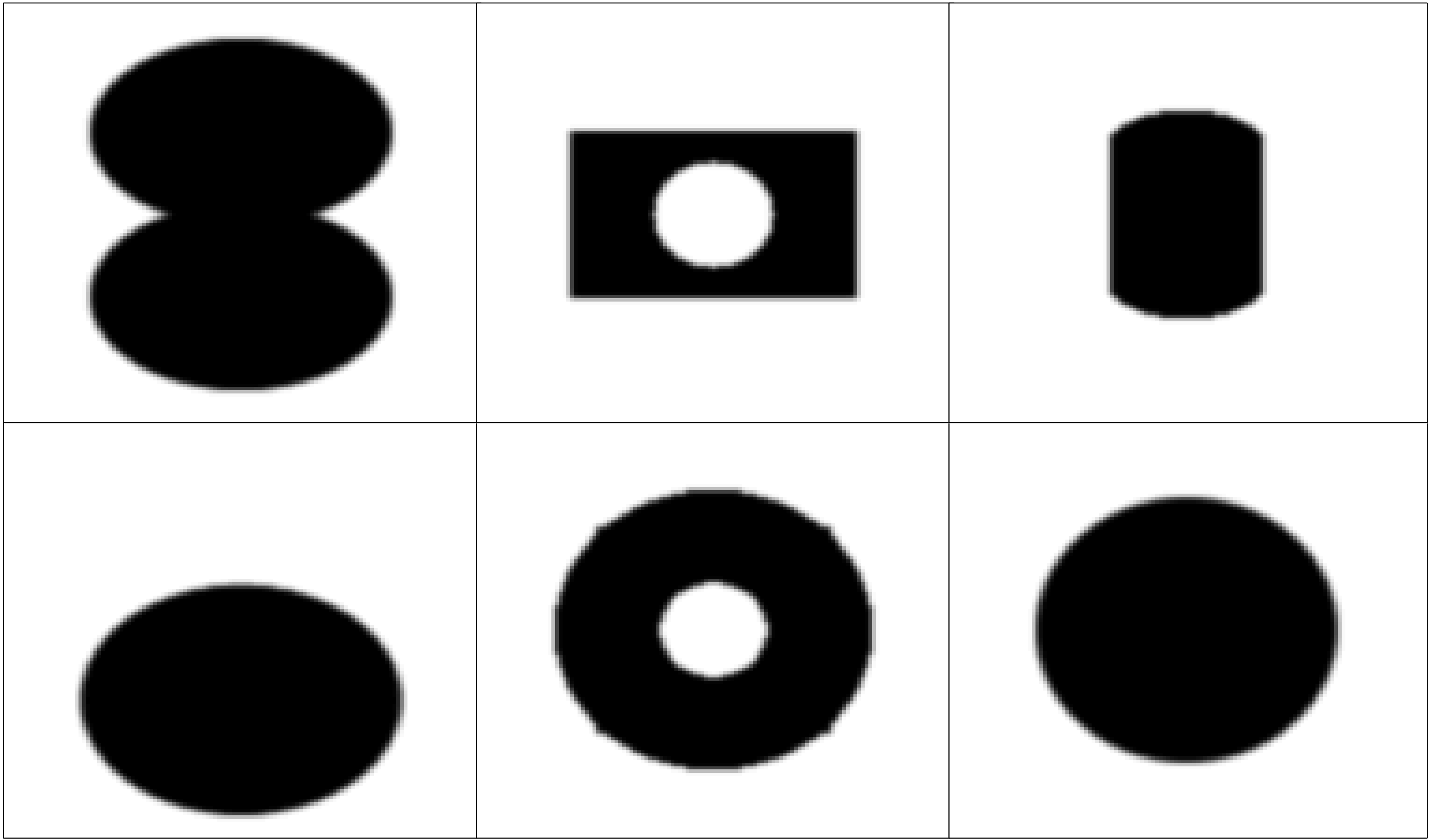



Template and it's gradients
Residuals
Momentum
Velocity
Deformation
Jacobian

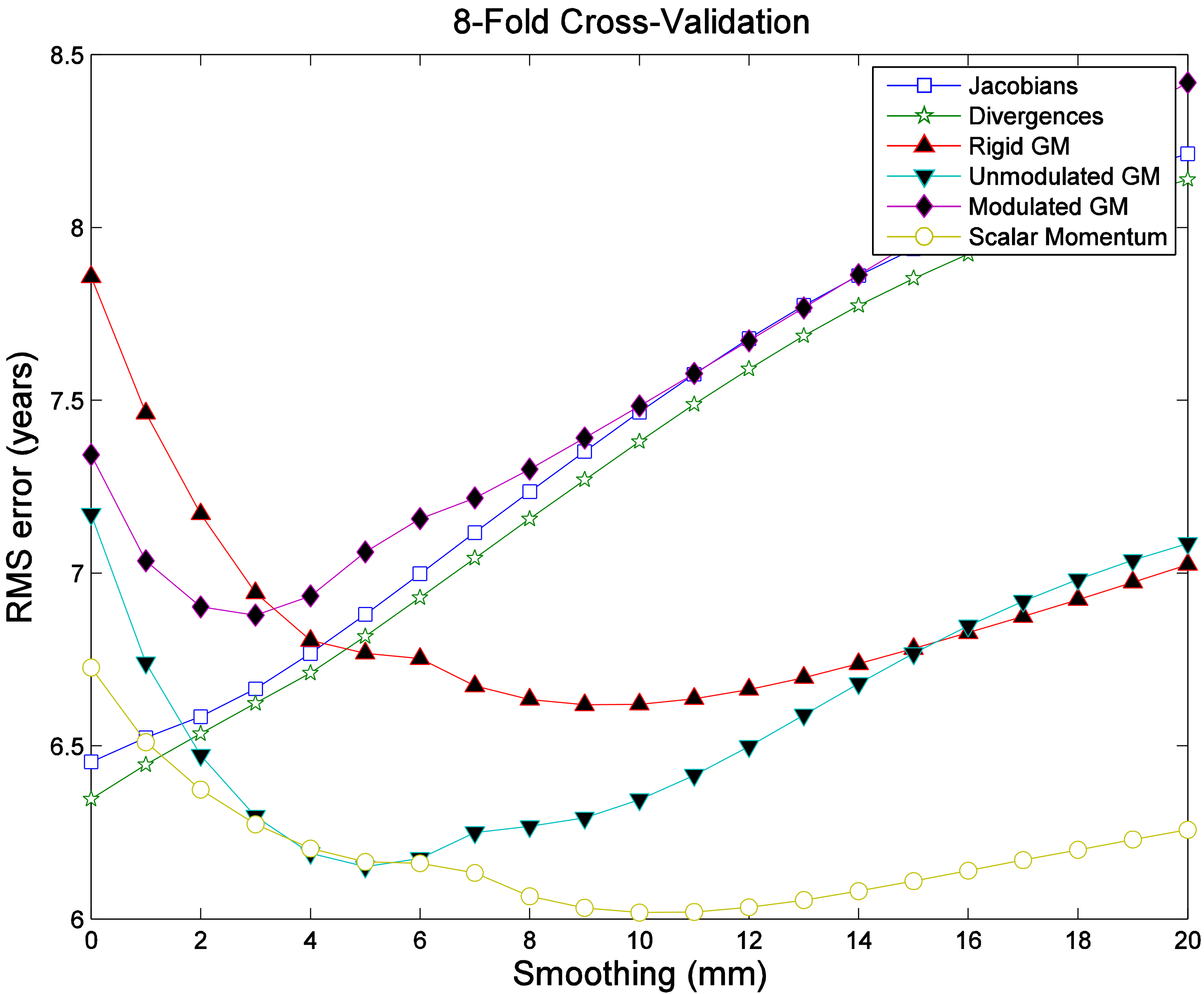
Reconstructed from Template & Residuals



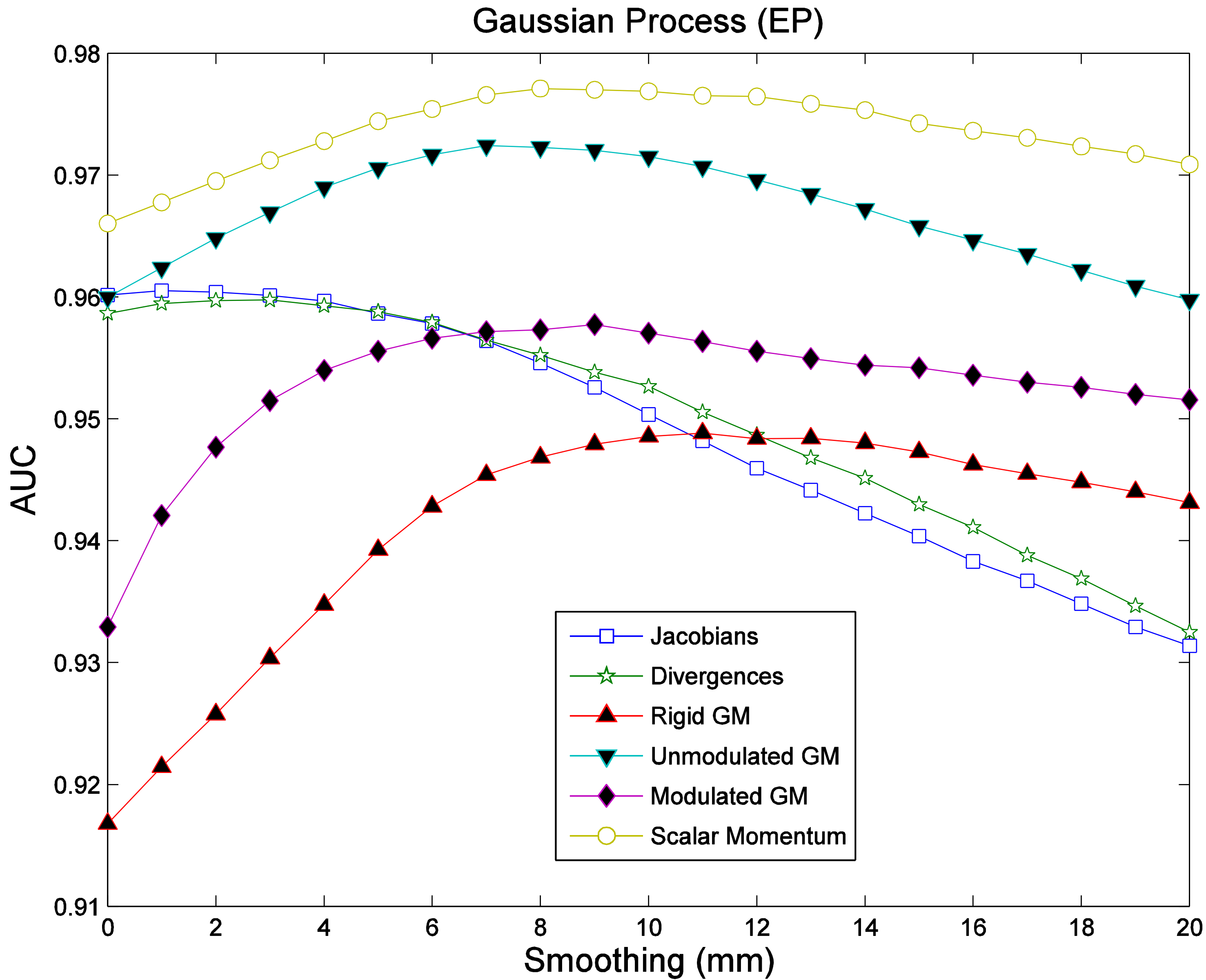
Original Examples



Predicting Age



Predicting Sex



MINING HOSPITAL IMAGES

Basic Research Data

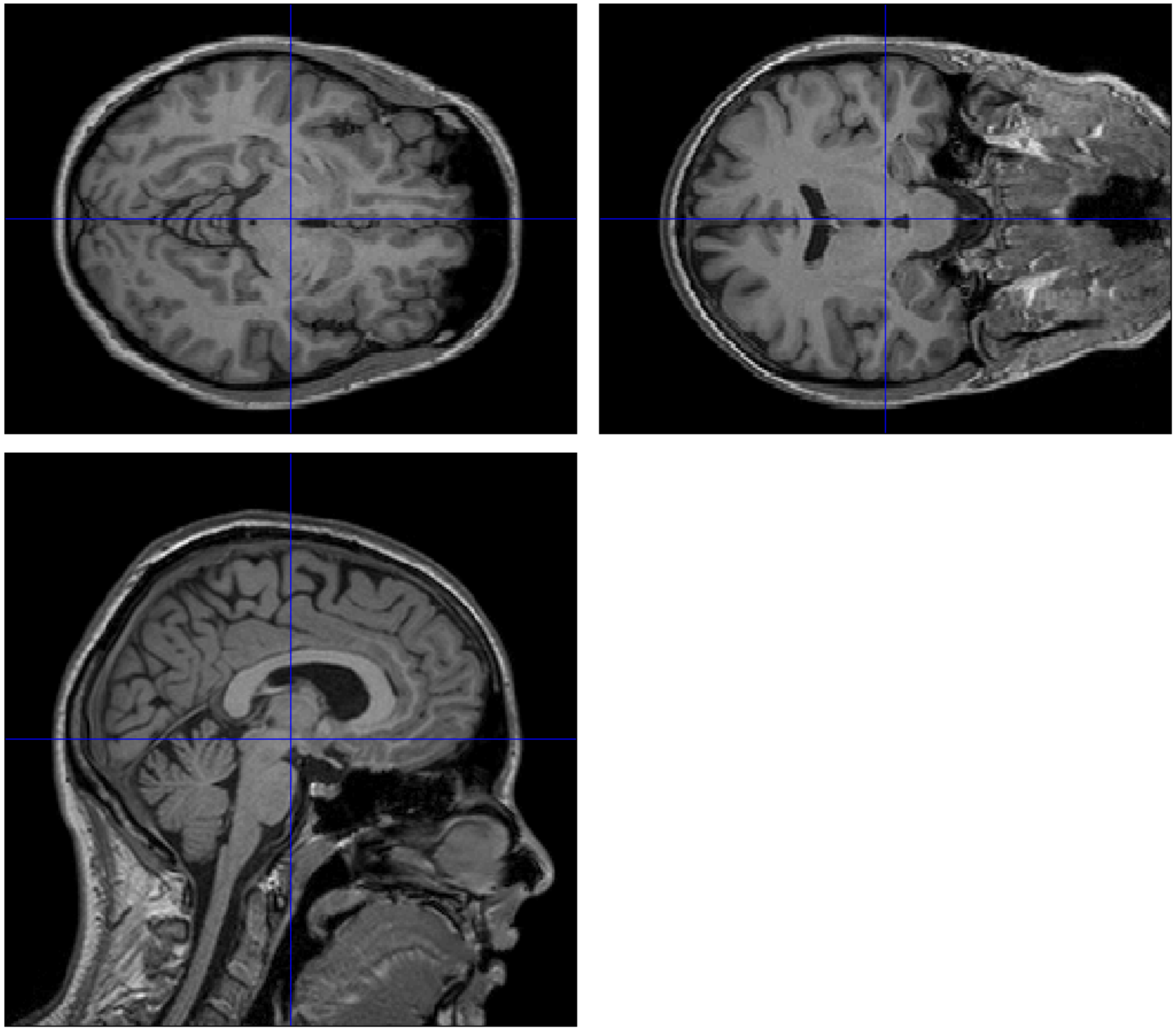
•Good quality images

•Well controlled experiments

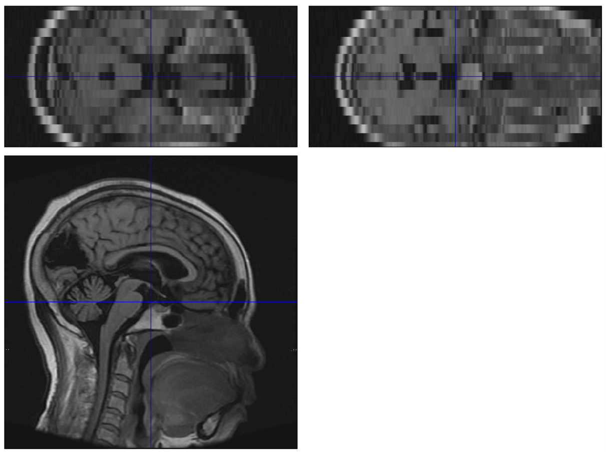
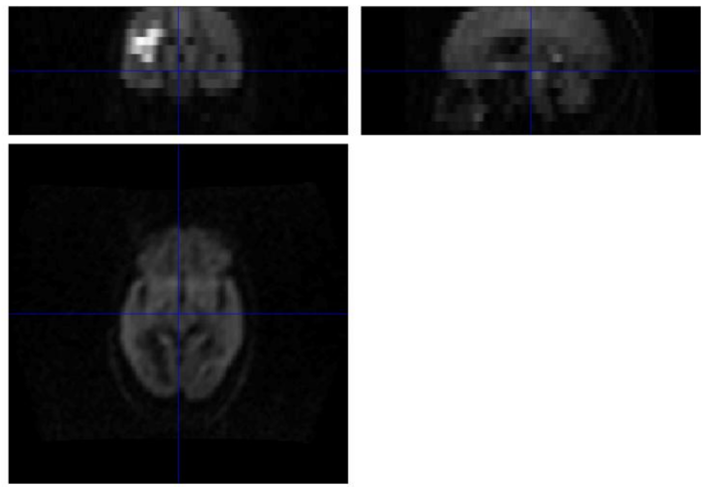
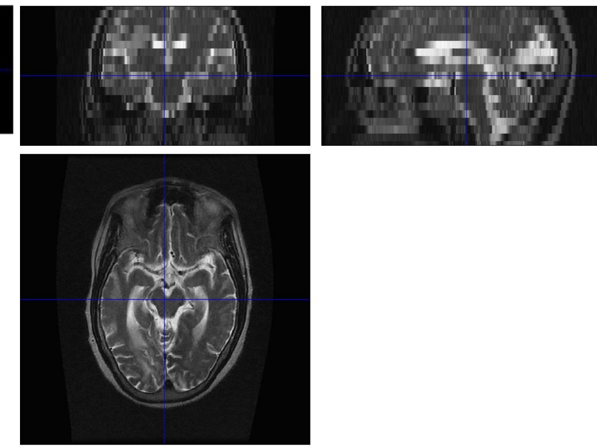
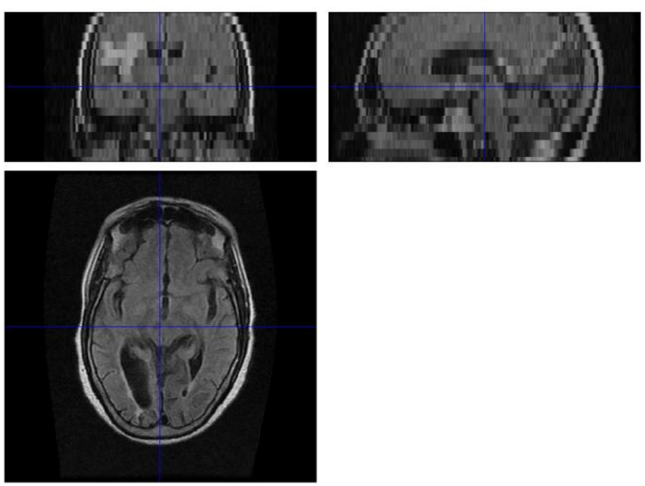
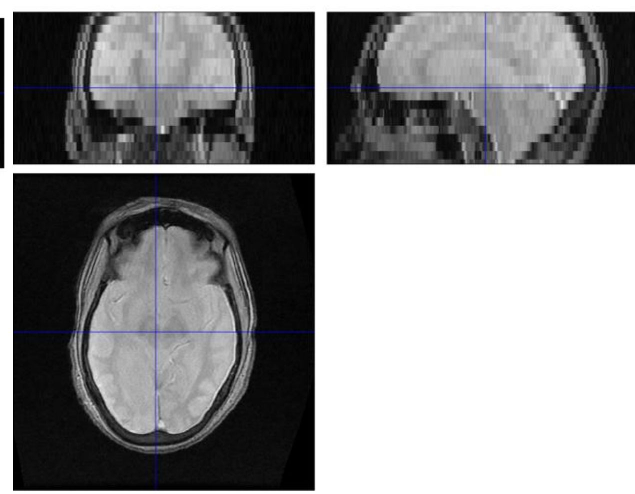
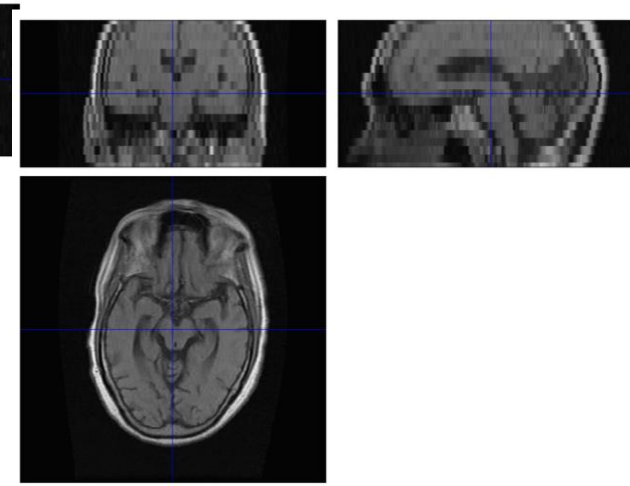
•Typically one scan per subject

•Mostly healthy subjects

•No major pathologies



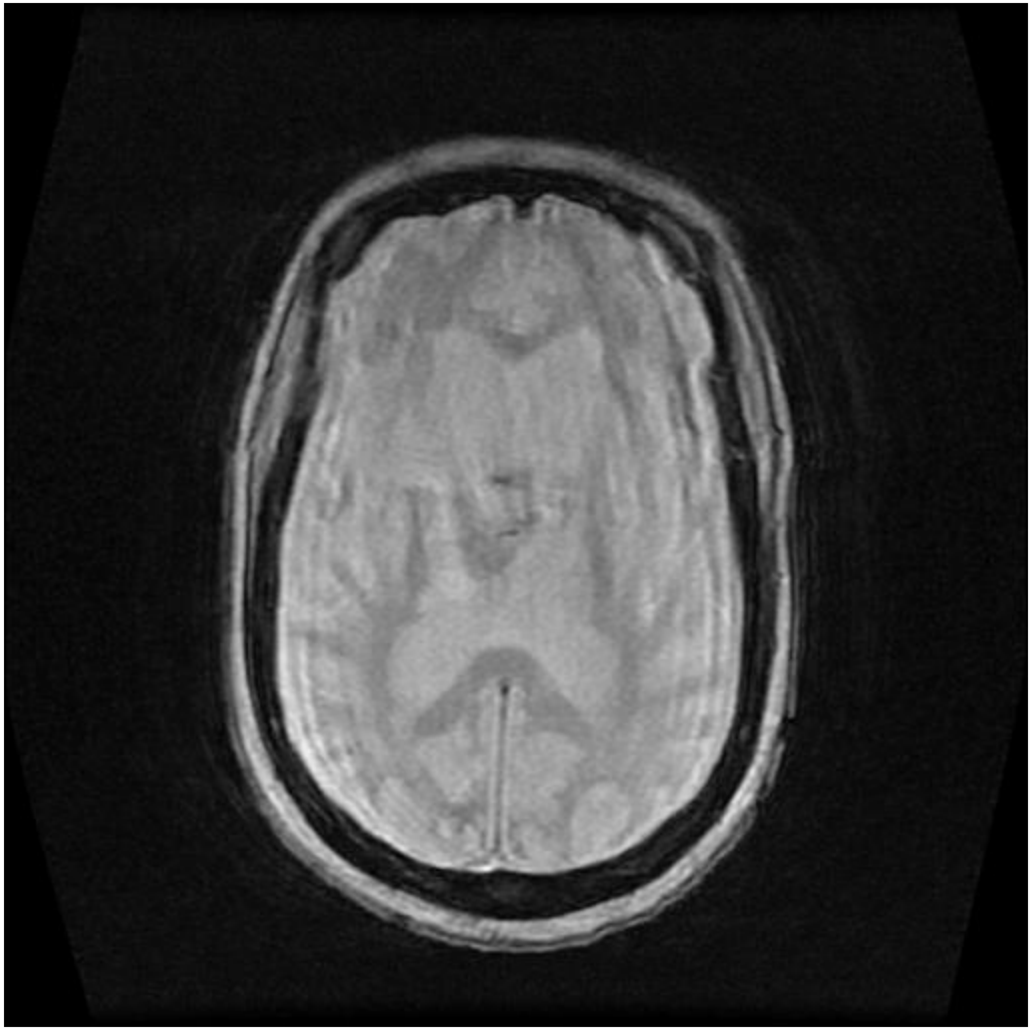
Hospital scans are a bit crappy

Movement Artifact

Patients move in the scanner.

Data is very hard to make use of.

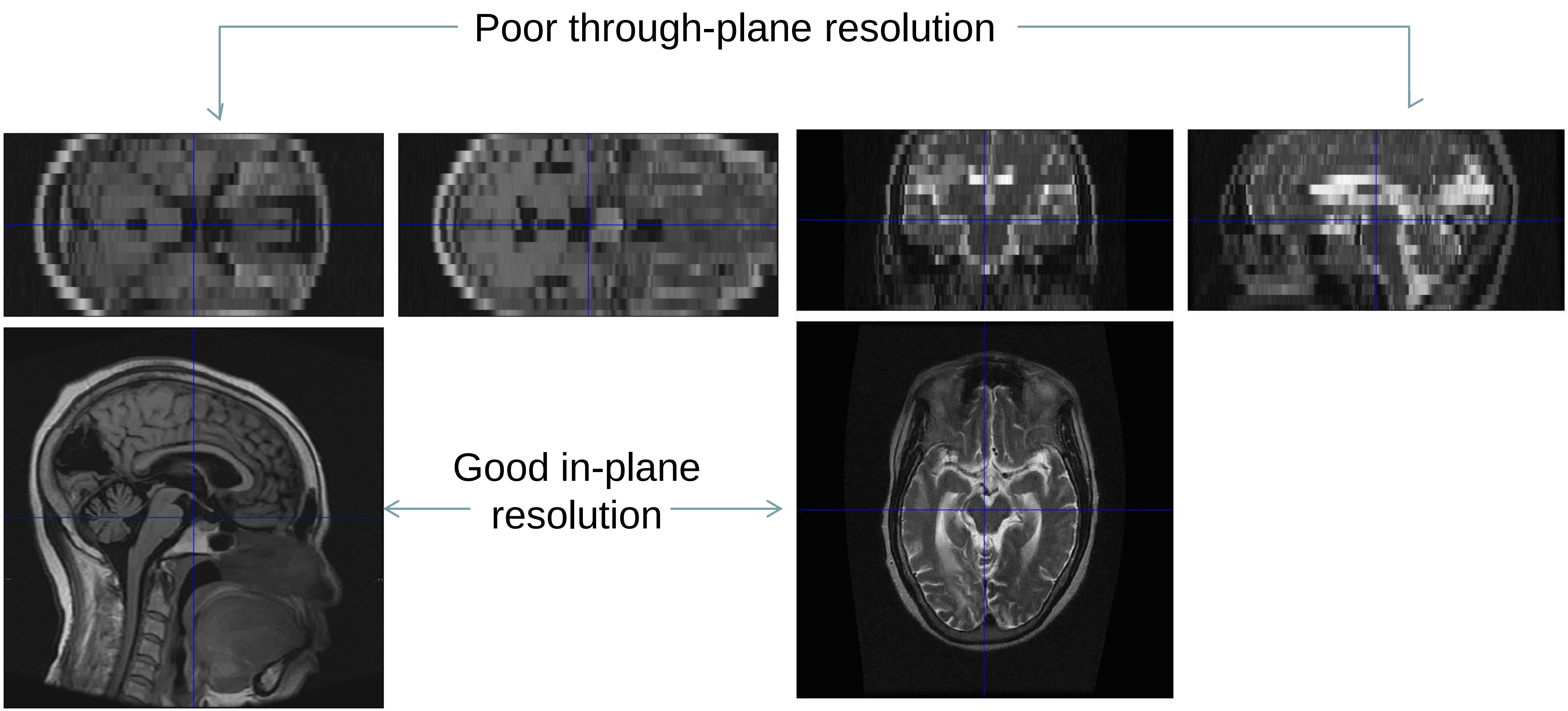


“Disease signature” of Parkinsons.

Few fully automated algorithms can make good use of hospital data.

- Most algorithms expect T1-weighted MRI with 1mm isotropic resolution.
- They do a poor job with hospital scans.
- More work is needed.

Same subject, different MRI contrasts, different image orientations.



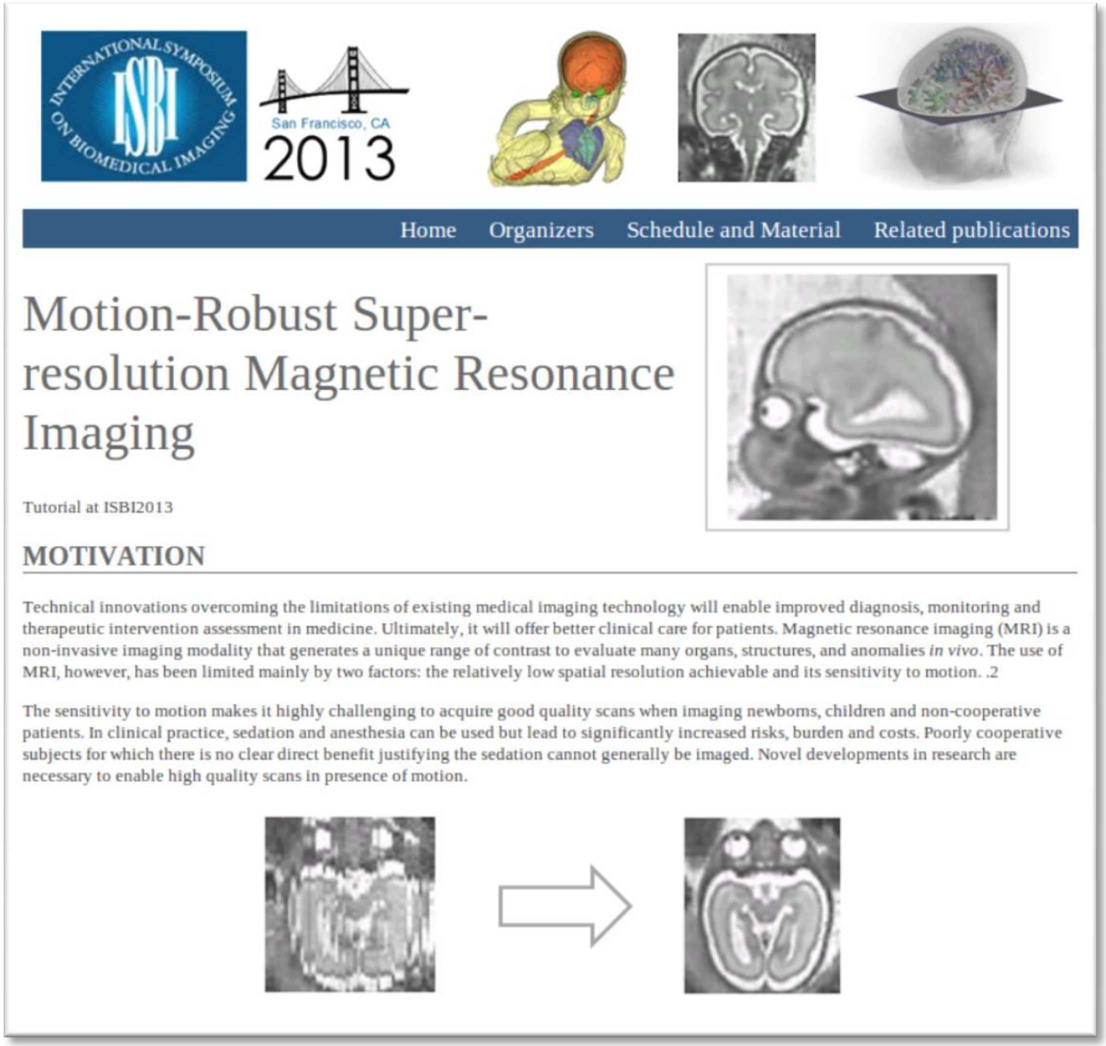
Recovering information

Several approaches for *super-resolution*.

Recover higher resolution signal from several low-resolution images taken from different views.

Work so far has assumed that all images are of the same modality.

Less straightforward for multi-modal data.



Generative Model for Resolution Recovery

Noise precision

Observed images

Mixing proportions

Means

$$p(\mathbf{F}, \mathbf{G}, \boldsymbol{\tau}, \mathbf{Z}, \boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\pi}) = \overbrace{\prod_{d=1}^{D} \left(p(\mathbf{f}_d|\mathbf{g}_d, \tau)p(\tau_d)\right)}^{\text{Probability of observations given the reconstructions}} \ldots$$

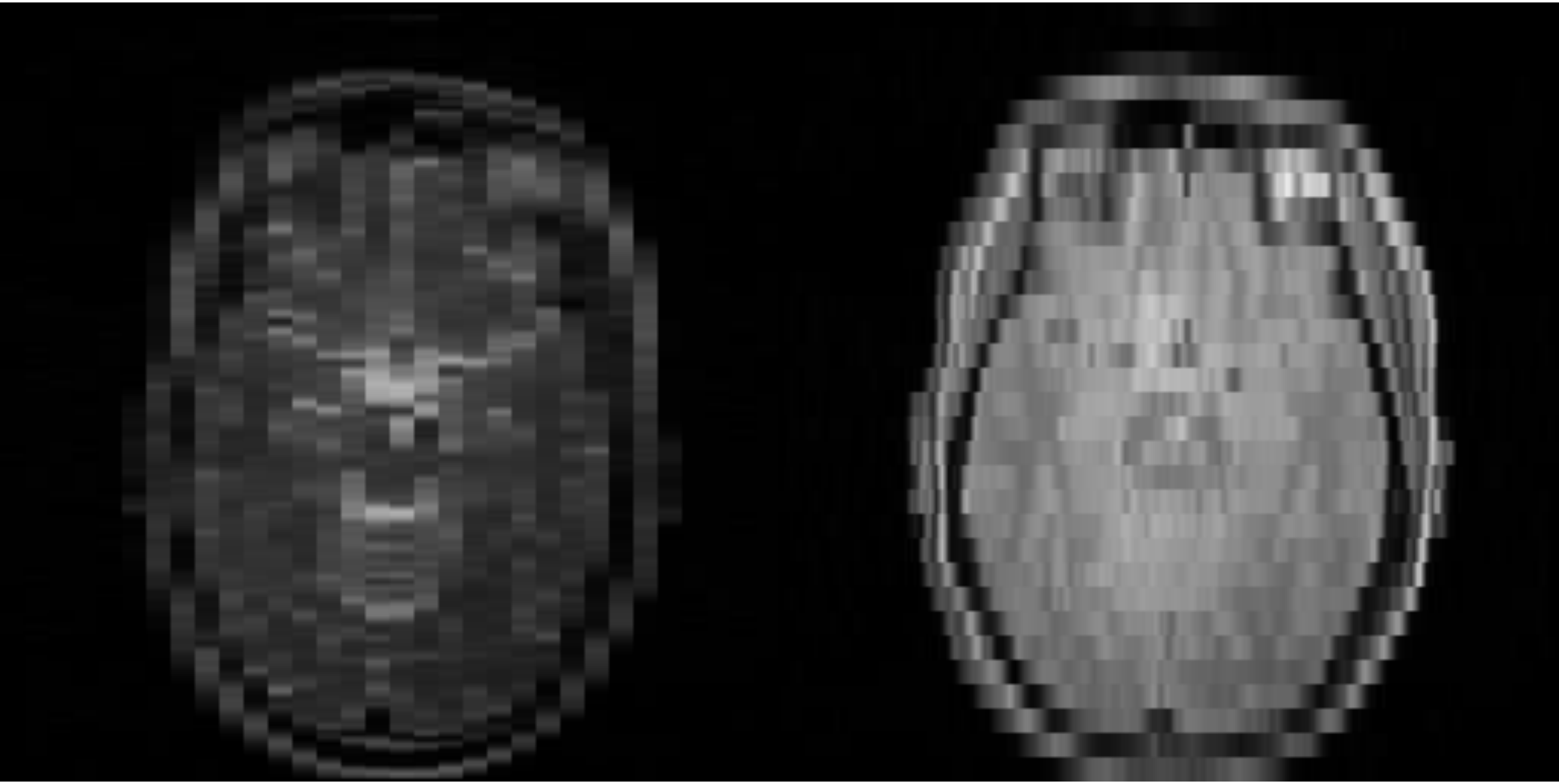
$$\underbrace{\prod_{n=1}^{N} \left(p(\mathbf{g}_{n,:}|\mathbf{z}_{n,:}, \boldsymbol{\mu}, \boldsymbol{\Lambda})P(\mathbf{z}_{n,:}|\boldsymbol{\pi})\right)}_{\text{Probability of reconstructions given the MoG parameters}} \overbrace{\prod_{k=1}^{K} \left(p(\boldsymbol{\mu}_k|\boldsymbol{\Lambda}_k)p(\boldsymbol{\Lambda}_k)p(\pi_k)\right)}^{\text{Prior probability of MoG parameters}}$$

Class labels

Reconstructed images

Precisions

Simple 2D simulations with 8mm thick “slices”



Prototype of resolution recovery attempt.