Australian National University

Geometric morphometric analysis

for quantifying and categorizing brain structure shape in diabetic and non-diabetic populations

Dr. Erin Walsh

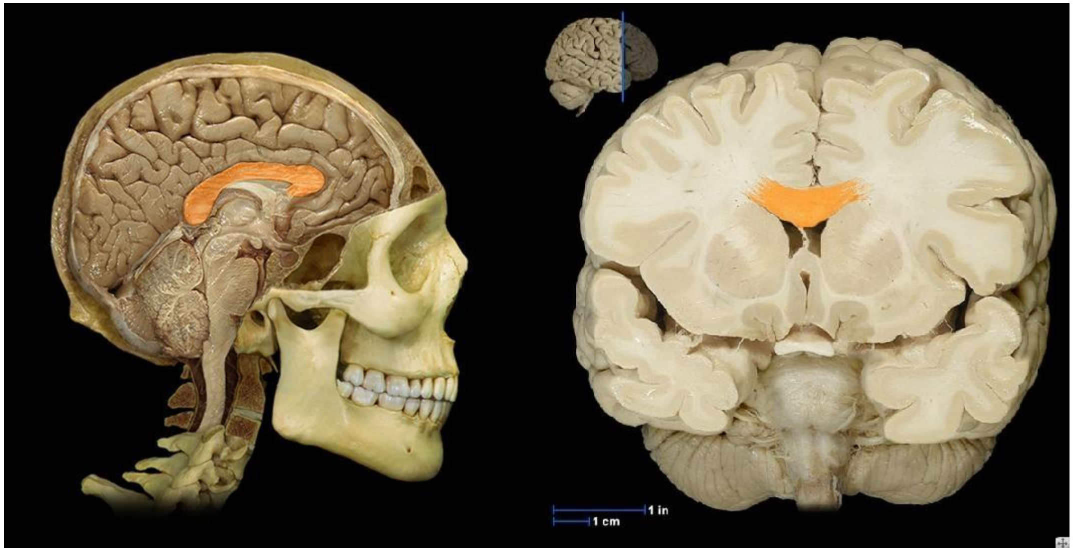
Dr. Marnie Shaw, Mark Fraiser, Kaarin Anstey & Nicolas Cherbuin

Centre for Research on Ageing, Health and Wellbeing

Research School of Population Health, ANU

Australian National University

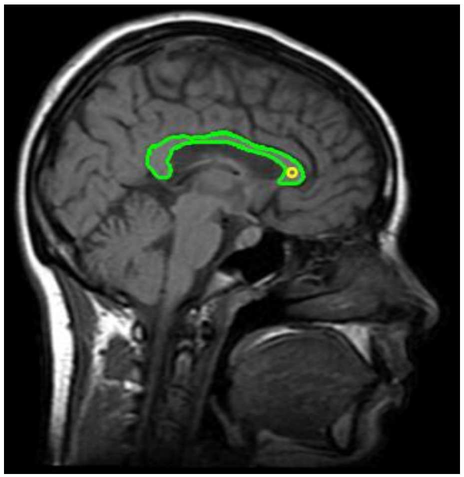
Shape can impact function

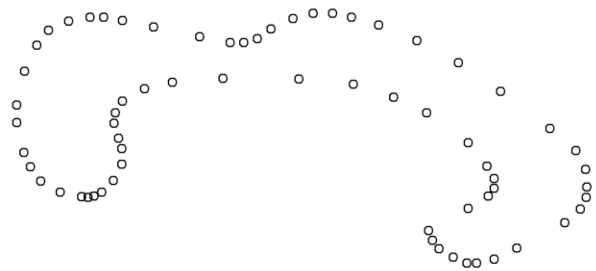


Sources: http://www.previninc.com/shop/media/detail/93433.jpg
http://www.matferbourgeatusa.com/content/images/thumbs/0001141_mise-en-bouche-mini-bent-spoon.jpeg
http://img.photobucket.com/albums/v120/tanalu/Spoon4.jpg
http://johnlewis.scene7.com/is/image/JohnLewis/231159253?$prod_exlrg$http://femininepeace.org/wp-content/uploads/corpus_callosum1337487504159.jpg

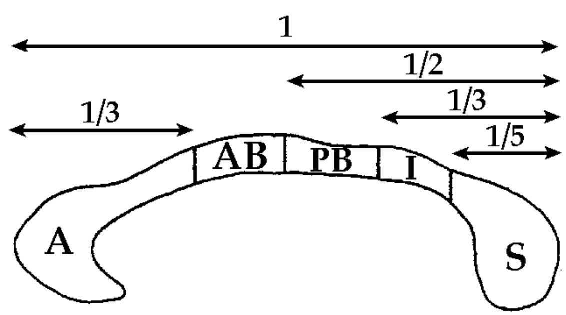
How to quantify shape?

A problem of dimension reduction





glm(Diabetes ~ area)



glm(Diabetes ~ areaA)
glm(Diabetes ~ areaAB)
glm(Diabetes ~ areaPB)

Geometric morphometric analysis

glm(Diabetes ~ shape)?

Source: http://cmp.felk.cvut.cz/cmp/demos/Medical/LevelSetSegmentation/images/corpusCallosum_seg.jpg

Australian National University

How to quantify shape?

Polygons and fouriers

Polygonisation

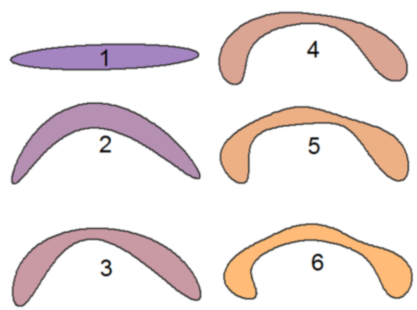
- Deconstruction of shape into polygons (here, rectangles)

Simplified coordinates

	x	y
1	1.499993	0.7500055
2	1.331718	0.8540184
3	1.163443	0.9391722
4	0.995168	0.9900272

Fourier

- Deconstruction of shape into harmonics



Harmonics

A1	A2	A3	B1	B2	B3
1	0.0384	-0.00559	5.39e-17	0.0758	-0.0197
1	0.0473	-0.07230	-2.67e-17	0.0674	-0.0275
1	0.0587	0.00156	-4.89e-17	0.0546	-0.0174
1	0.0714	-0.03920	-5.25e-17	0.0887	-0.0357
1	0.0248	0.02420	0.00e+00	0.0522	-0.0111
1	0.0277	0.02920	1.62e-16	0.0558	-0.0317

eFourier

tFourier, rFourier

Sources: http://soccernewsday.com/userfiles/images/02-17-2012.jpg
http://f.tqn.com/y/quotations/1/W/h/D/173341886.jpg

How to quantify shape?

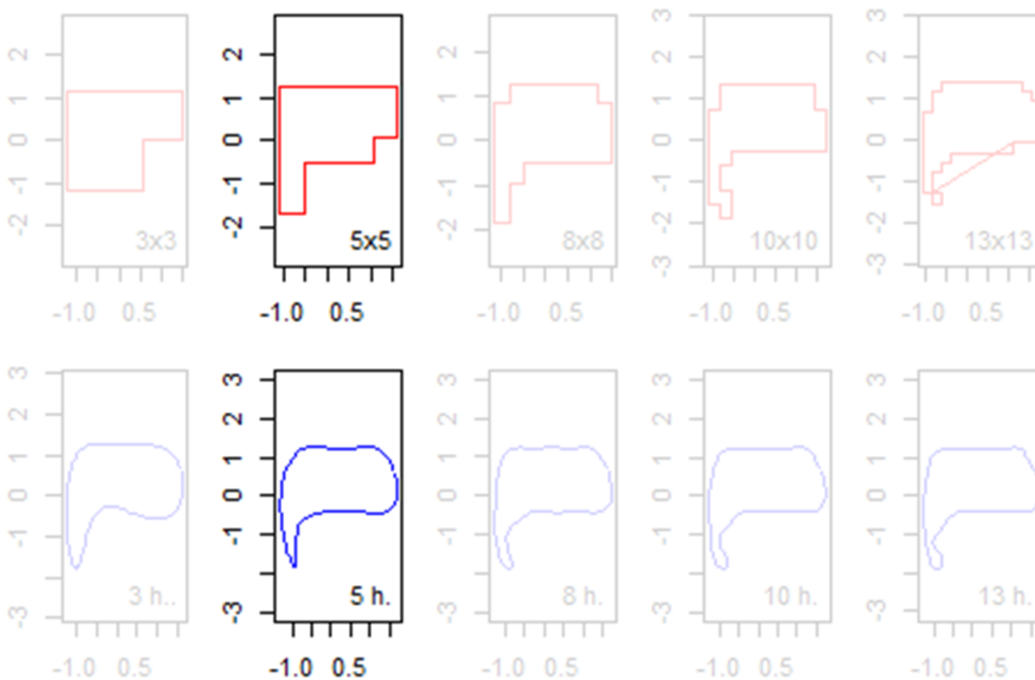
Polygons and fouriers

To get a little, you need to give a little…

Polygons
Variable dimension reduction
- 3x3 grid -> 6-9 coordinates
- 13x13 grid -> 6-169 coordinates

eFourier
Consistent dimension reduction
- 3 harmonics -> 6 harmonic pairs
- 13 harmonics -> 26 harmonic pairs

For this talk, we're going with **5**
- 5x5 grid -> 6-25 coordinates
- 5 harmonics -> 10 harmonic pairs



So, what works best?

… Wait, what is *'best'*?

What I've got:
How well the different methods can recover shapes

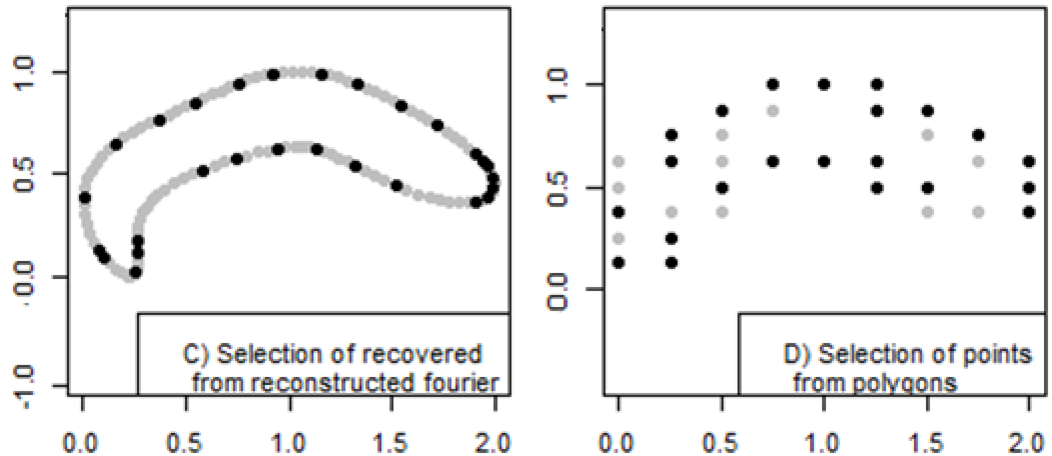
What I also want:
How sensitive the methods are to shape differences when plugged into subsequent (epidemiologist friendly) analyses

Source: http://www.theblaze.com/wp-content/uploads/2013/08/600x4503.jpg
http://georgiamidwife.org/wp-content/uploads/2015/07/We-want-you-image.png

Metrics for evaluation

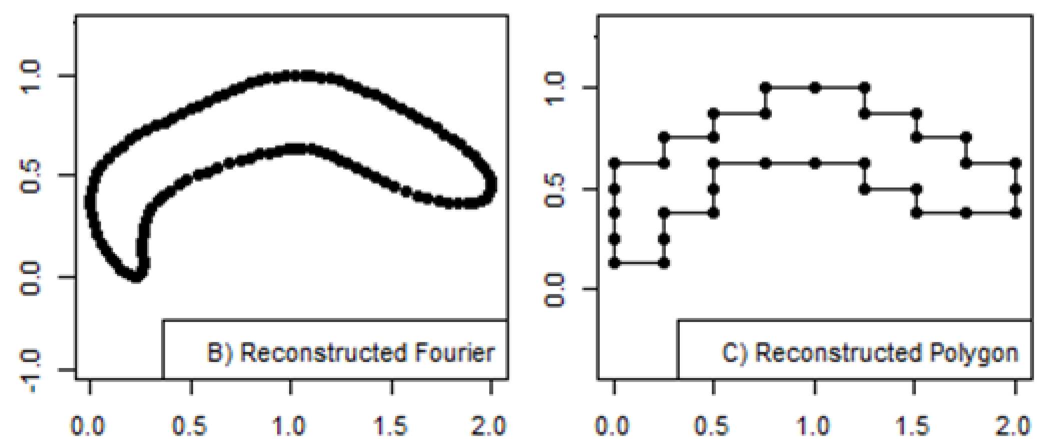
$$\frac{\sum_{k=1}^{K}\sum_{j=1}^{J}|Original_{kj} - Reconstructed_{kj}|}{K}$$

Outline fidelity



$$|Original_a - Reconstructed_a|$$

Area fidelity



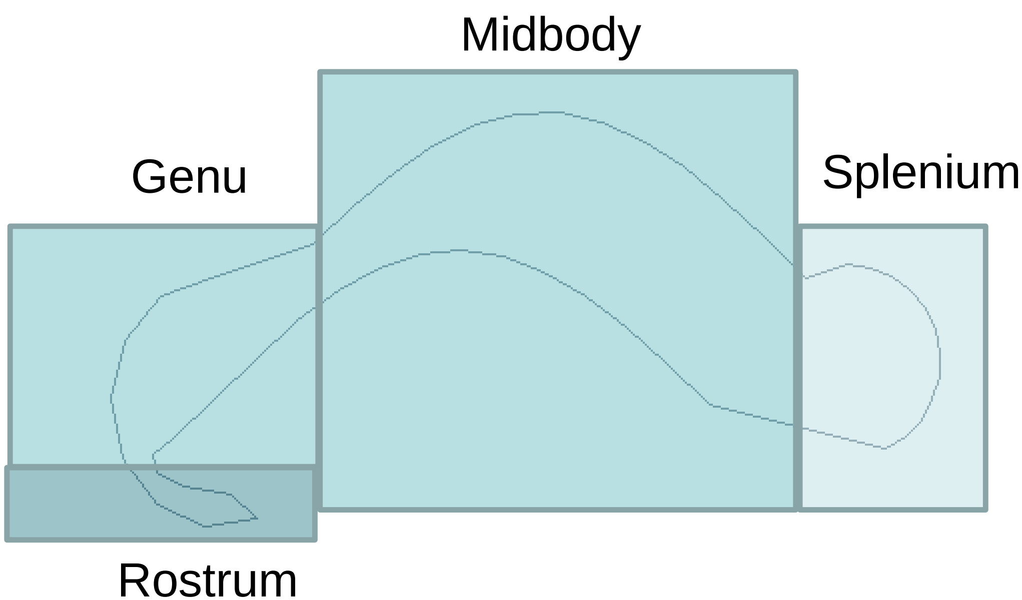
Shapes for evaluation

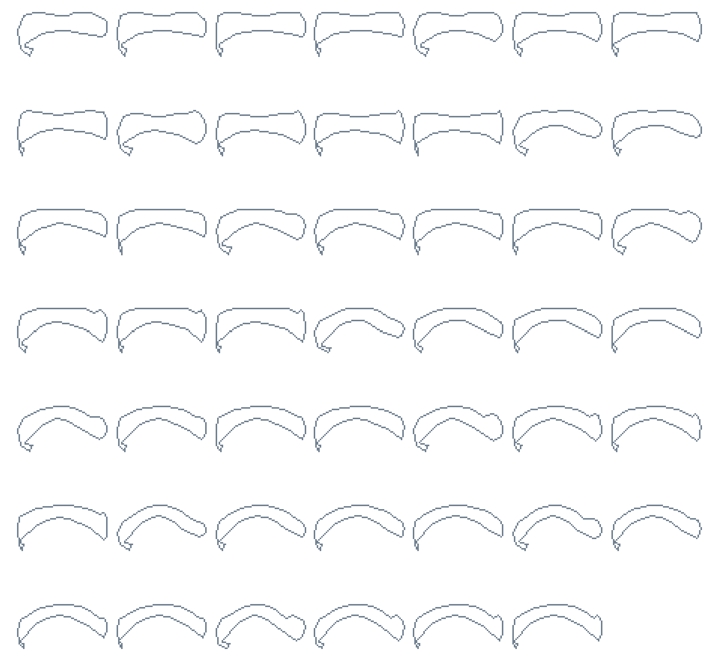
A lot of variability in ‘real’ data… so we’ll start with some fake data.

- Systematically modifiable test shapes with similar elements to the corpus callosum

Caveat: WITHOUT using tools we will then use to describe the shapes

- – Solution: Frankenstein’s monster of circles and sine waves:

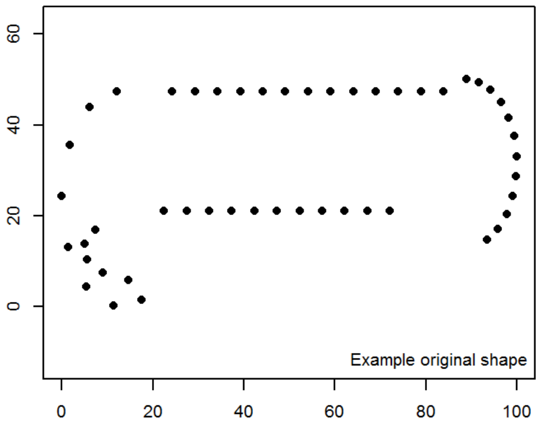




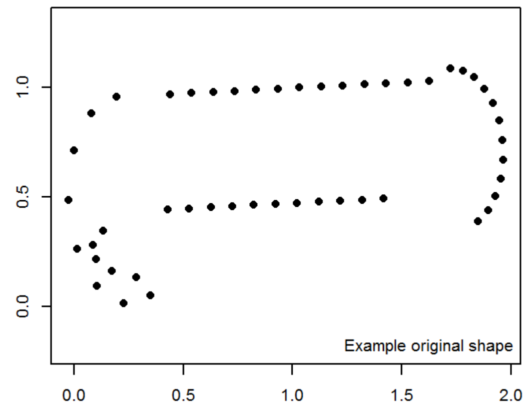
Modification of things that shouldn't change…

Non-shape differences

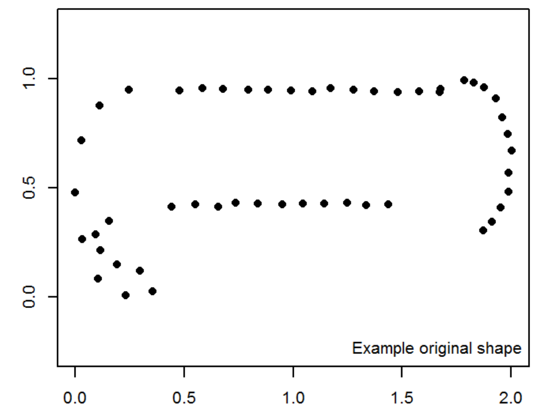
Scale



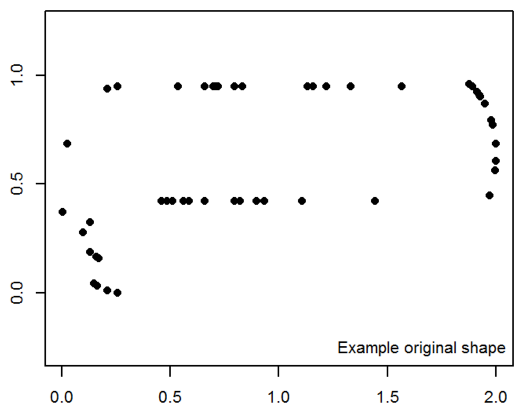
Orientation



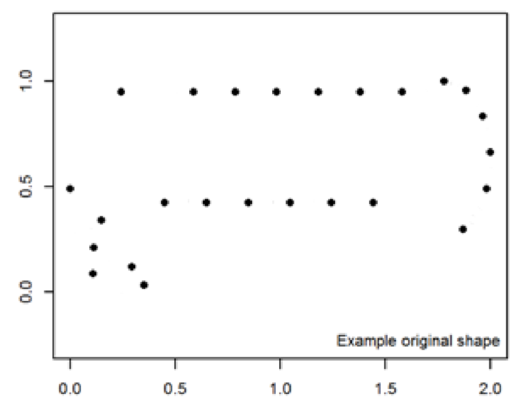
Jitter in outline



Randomness in outline



Number of points

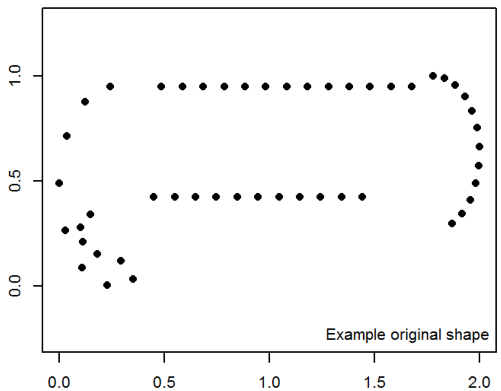


Australian National University

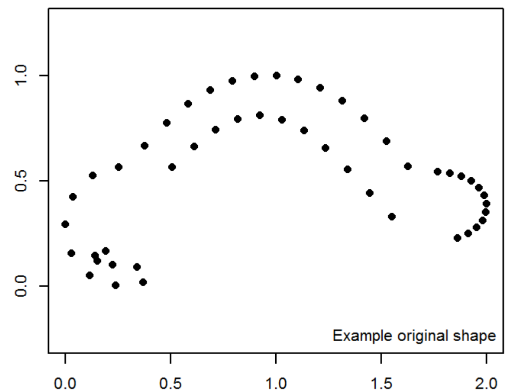
Modification of things that should change….

Real shape differences

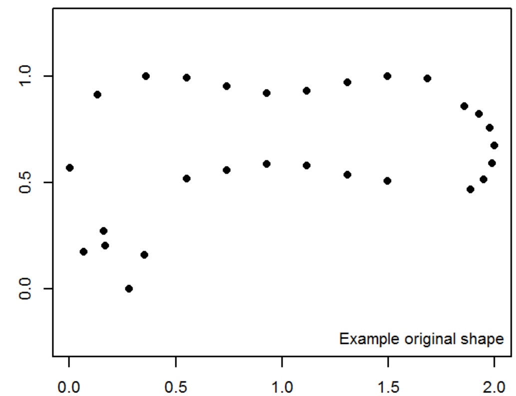
Length



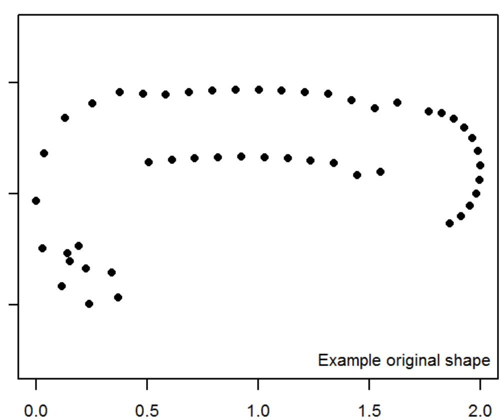
Curve



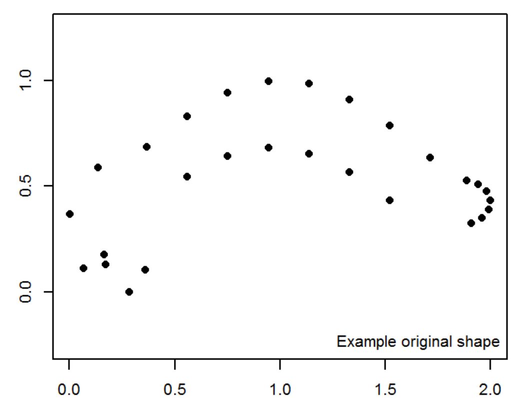
Regularity



Splenium size

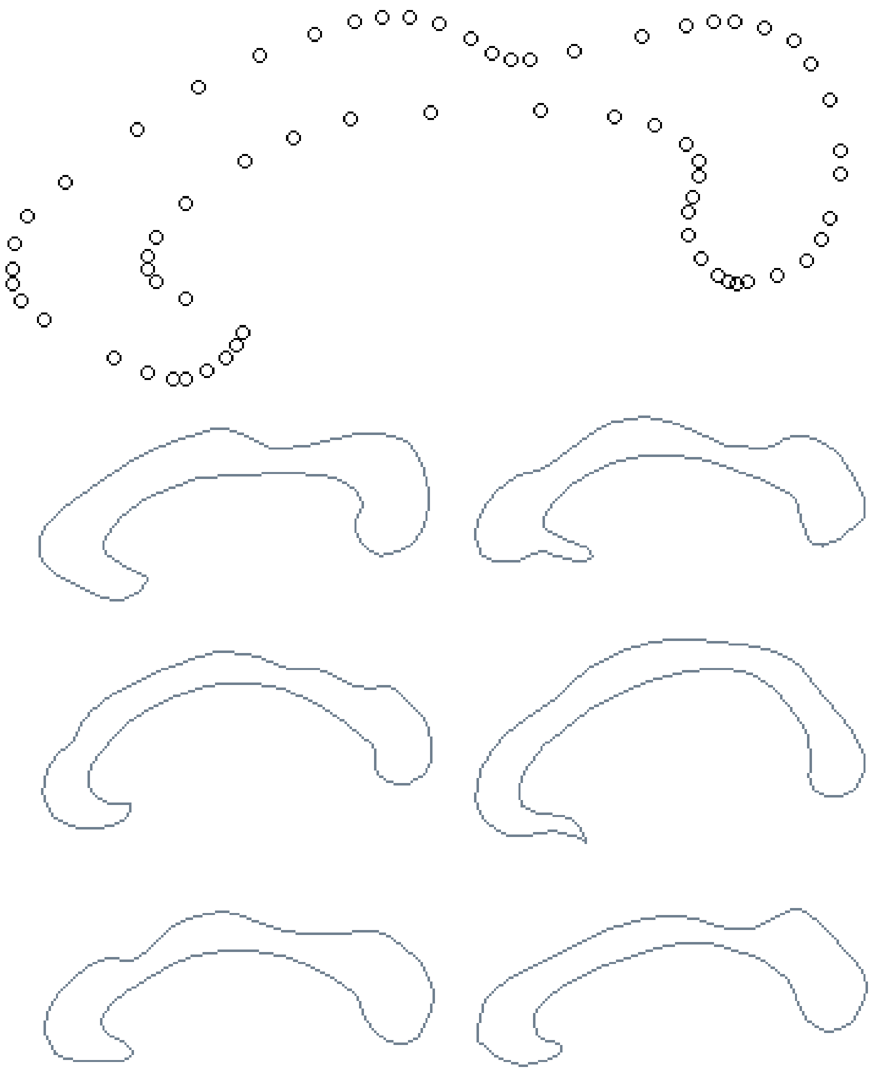


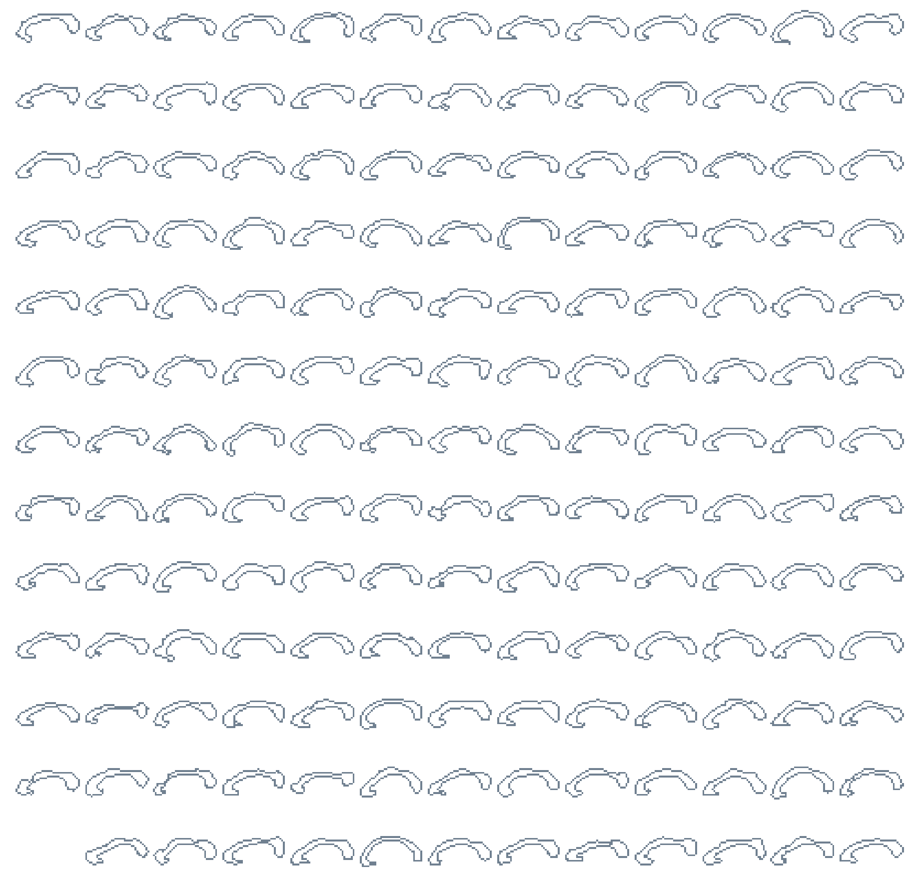
Everything changing!



Australian National University

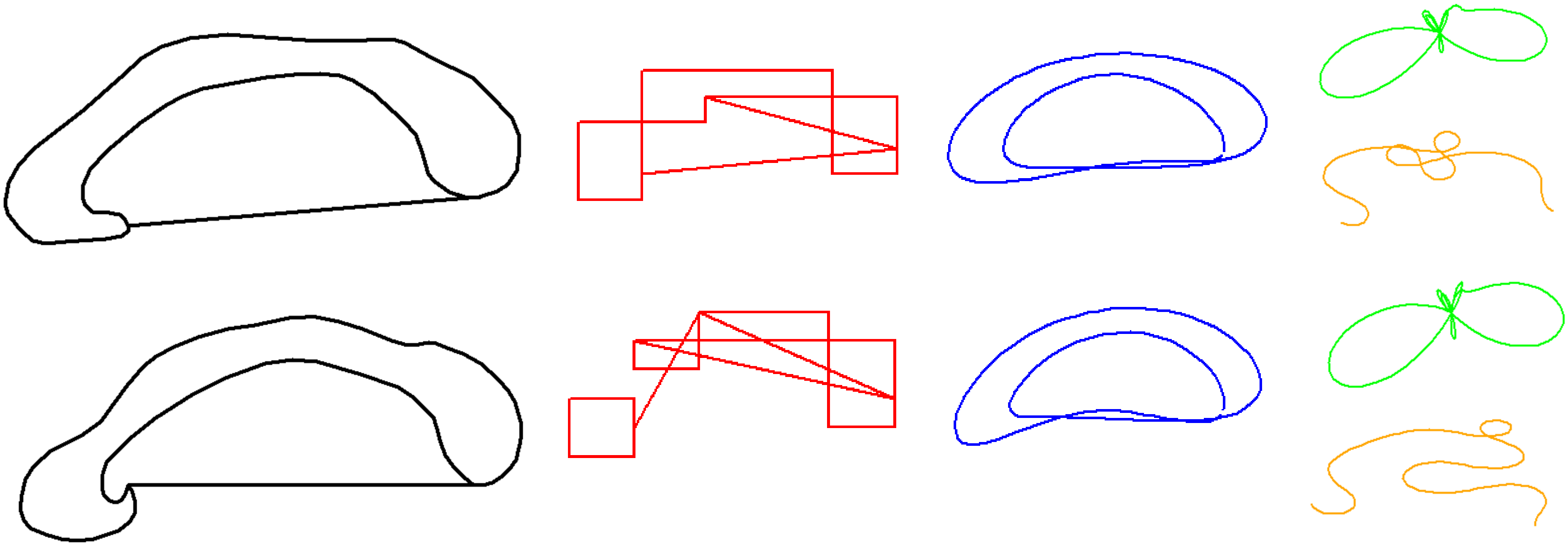
… then try it with the real thing.



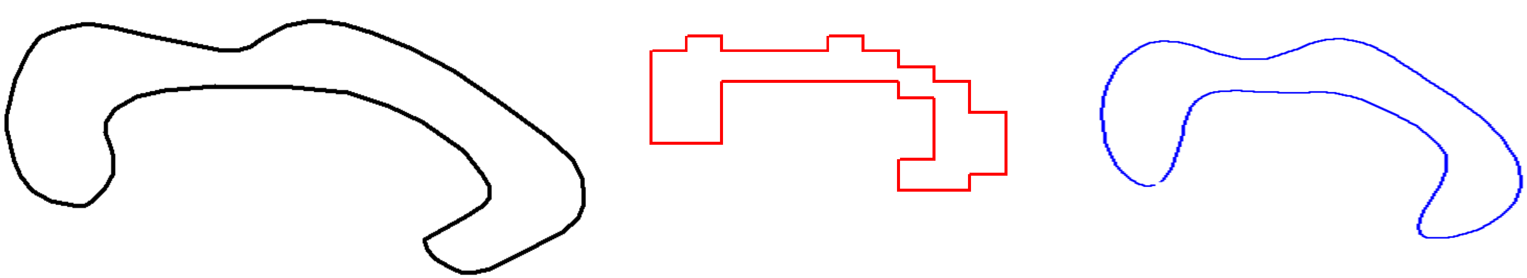


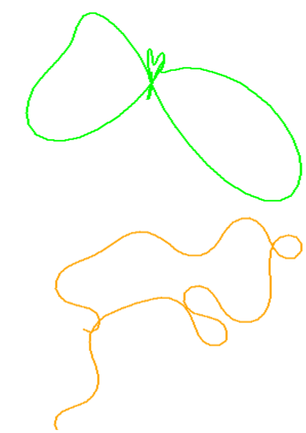
168 traces
Community-living adults from the PATH Study.

The eyeball test



… Ramp up the detail + clean up input shape

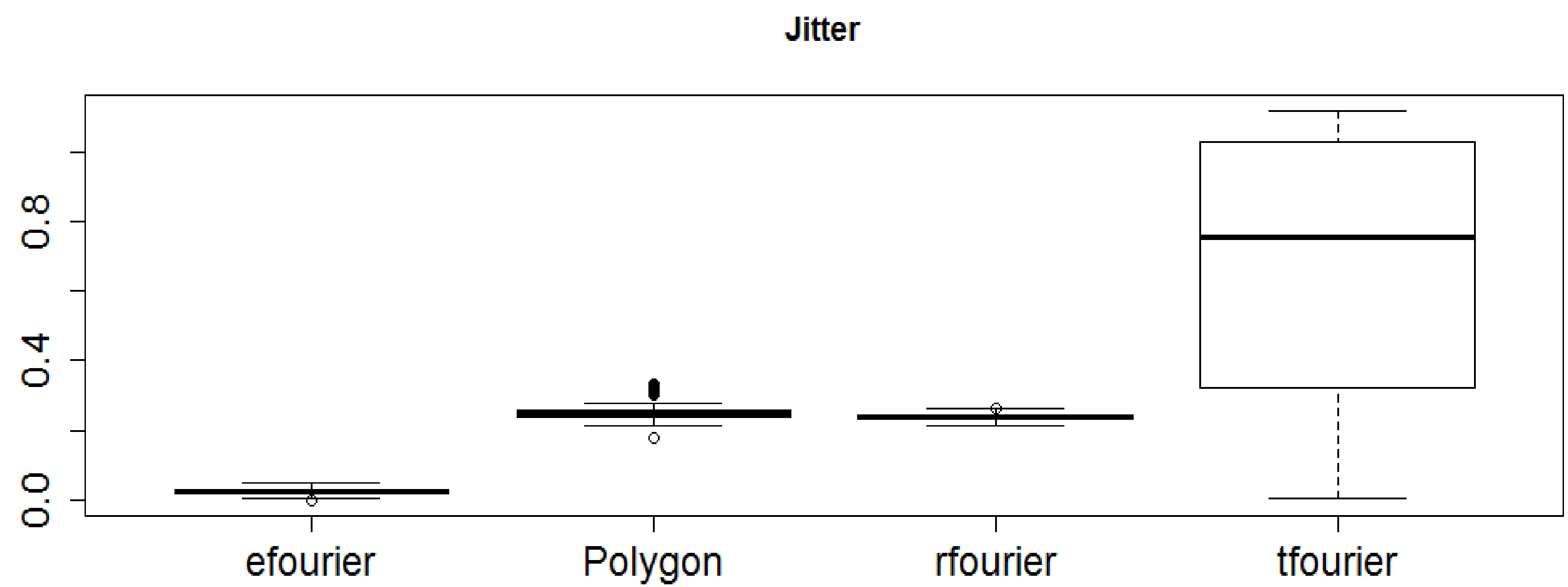




Quantifying ‘best’

Fidelity of area reconstruction $|Original_a - Reconstructed_a|$

glm(area difference ~ method) Smaller-is-better	(Intercept)	Polygon	rFourier	tFourier
Multiple parameters	0.04 [0.02,0.05]	0.6 [0.58,0.62]	0.29 [0.28,0.31]	0.21 [0.2,0.23]
Real data				
Real Shapes	[illegible] [196.18,309.07]	[illegible] [755.74,915.4]	1065 [985.17,1144.83]	1067.75 [987.92,1147.59]



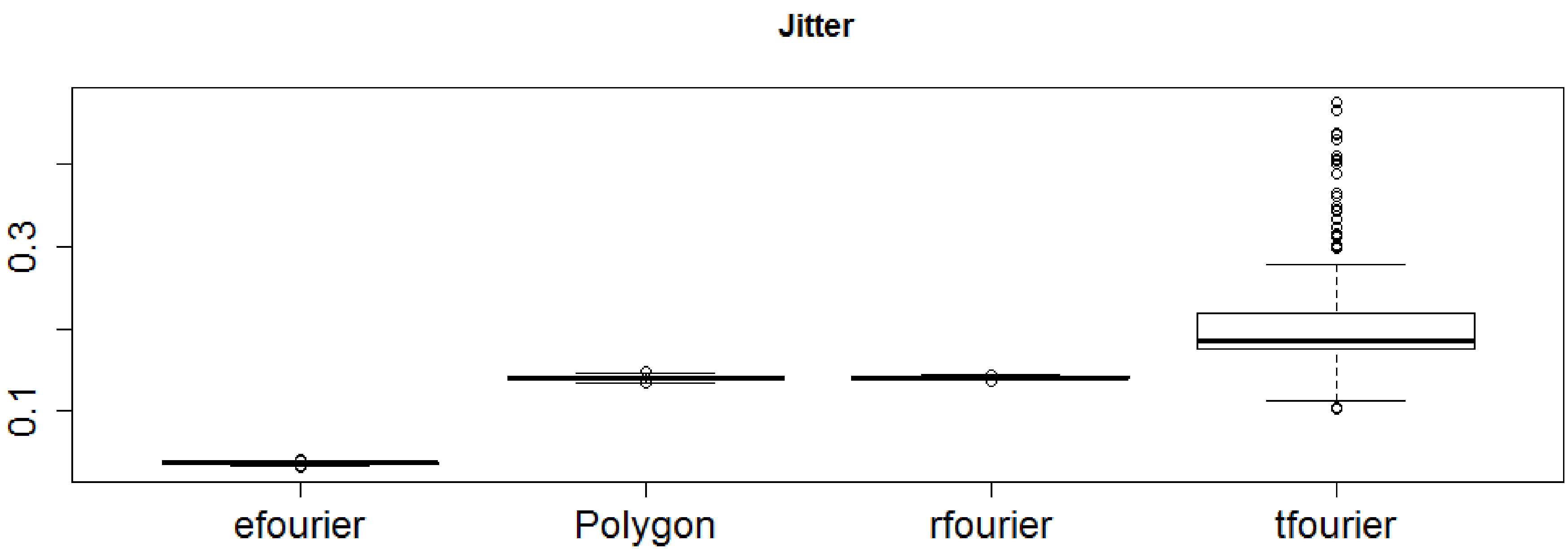
Interpretation: everything is worse than e-fourier

Base group: eFourier. Harmonics: 5, grid: 5x5. n=1000, except for real shapes (n=168). [denotes 95% CI.

Quantifying ‘best’

Fidelity of outline reconstruction $$\frac{\sum_{k=1}^{K}\sum_{j=1}^{J}|Original_{kj} - Reconstructed_{kj}|}{K}$$

glm(outline difference ~ method) Smaller-is-better	(Intercept)	Polygon	rFourier	tFourier
Real data				
Real Shapes	[illegible]	[illegible]	[illegible] [2.28,2.51]	3.37 [3.25,3.48]

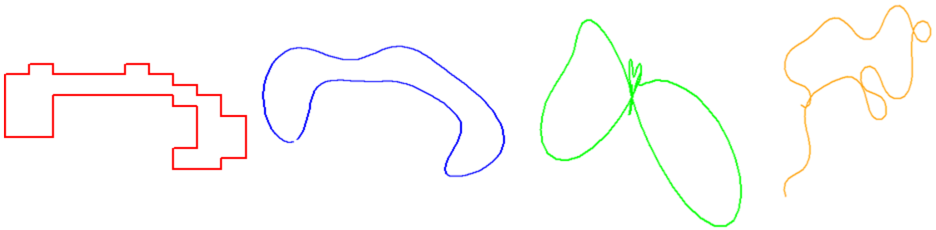


Interpretation: everything is worse than e-fourier

Base group: eFourier. Harmonics: 5, grid: 5x5. n=1000, except for real shapes (n=168). [denotes 95% CI.

All signs point to…

- Fidelity of outline reconstruction
- Fidelity of area reconstruction



Quantifying ‘best’

Q: How to quantify sensitivity to shape differences?

e.g. If I systematically modify the size of the splenium, at what point does the method register the difference?

(binary and/or continuous outcomes)

Source:
http://img2.wikia.nocookie.net/__cb20121213183536/es.pokemon/images/3/38/EP768_Leavanny_debilitado.png

The goal:

Quantify and categorize corpus callosum shape in diabetic and non-diabetic populations.

The question:

How best to quantify shape?

The (provisional) answer:

eFourier analysis

The next step:

- Better evaluation *sensitivity* to shape
  (Q: If I systematically modify the size of the splenium, at what point does the method register the difference?)
- eFourier frolic
  (Q: any interesting suggestions beyond MANOVA and PCA?)

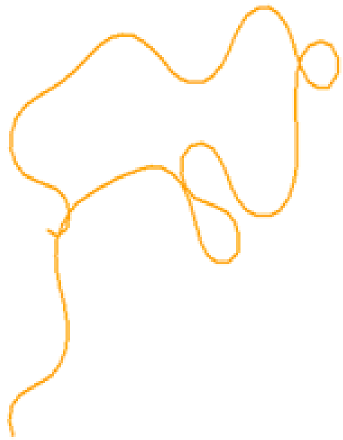
Australian National University

Acknowledgements

Thanks to:

- N. Cherbuin, M. Shaw, M. Fraser, the PATH team, interviewers and participants (for the data)
- V. Bonhomme (for the Fourier analysis r package that made all this practical)
- The Australian Research Council, Australian National Health and Medical Research Council, and National Infrastructure Facility (for the funding)

Happy to share R spaghetti code underlying this talk!



Bonus conclusion: rFourier = spaghetti?

We want YOU to suggest…

- Better evaluation *sensitivity* to shape
  (Q: If I systematically modify the size of the splenium, at what point does the method register the difference?)
- eFourier frolic
  (Q: any interesting suggestions beyond MANOVA and PCA?)

Stuff that didn’t fit in the presentation

Things to do

- (intended) full list of comparisons:
  - Constructed
  - Real (manually traced)
  - Real (automatically traced)
- Other similarity tests
  - Presence/absence of vectorized pixle in raster space (typically used in neuro for inter-rater reliability for area traces)
    - Jaccard index
    - Sørensen–Dice coefficient
  - ICC instead of glm

- Possible sensitivity tests
  - Sample output from 1000 shapes, 100 per group, increase distance between groups, GLM until significance
  - Simulate binary or continuous correlates for shapes, see how big the correlate has to be before GLM picks out a predictive relationship

- Poor tFourier and rFourier
  - Re-run with re-sampled contours so their assumptions are met so they have a fighting (if unrealistic) chance

Australian National University

Numeric results

Fidelity of area reconstruction

$$|Original_a - Reconstructed_a|$$

glm(area difference ~ method) Smaller-is-better	(Intercept)	Polygon	rFourier	tFourier
Things which shouldn't change				
Scale	7637.48 [-2663.31,17938.27]	73587.84 [59020.33,88155.36]	67581.88 [53014.36,82149.39]	347059.85 [332492.34,361627.37]
Orientation	0.03 [0.02,0.05]	1.07 [1.05,1.08]	0.1 [0.08,0.11]	1 [0.99,1.02]
Jitter in outline	0.02 [0.01,0.03]	0.22 [0.21,0.24]	0.21 [0.2,0.23]	0.67 [0.66,0.69]
Randomness in outline	0.04 [0.02,0.06]	0.51 [0.48,0.53]	0.2 [0.18,0.23]	0.77 [0.74,0.79]
Number of points	0.03 [0.02,0.05]	1.07 [1.05,1.08]	0.1 [0.08,0.11]	1 [0.99,1.02]
Things which should change				
Midbody length	0.04 [0.03,0.05]	0.75 [0.73,0.76]	0 [-0.02,0.01]	0.58 [0.56,0.59]
Midbody curve	0.01 [0,0.03]	0.59 [0.57,0.61]	0.34 [0.32,0.36]	0.09 [0.07,0.11]
Midbody regularity	0.01 [0,0.02]	1.11 [1.1,1.12]	0.16 [0.15,0.17]	0.41 [0.4,0.42]
Splenium size	0.11 [0.09,0.12]	0.93 [0.91,0.95]	0 [-0.02,0.02]	0.52 [0.5,0.54]
Multiple parameters	0.04 [0.02,0.05]	0.6 [0.58,0.62]	0.29 [0.28,0.31]	0.21 [0.2,0.23]
Real data				
Real Shapes	252.62 [196.18,309.07]	835.57 [755.74,915.4]	1065 [985.17,1144.83]	1067.75 [987.92,1147.59]

Base group: eFourier. Harmonics: 5, grid: 5x5. n=1000, except for real shapes (n=168). [denotes 95% CI.

Australian National University

Numeric results

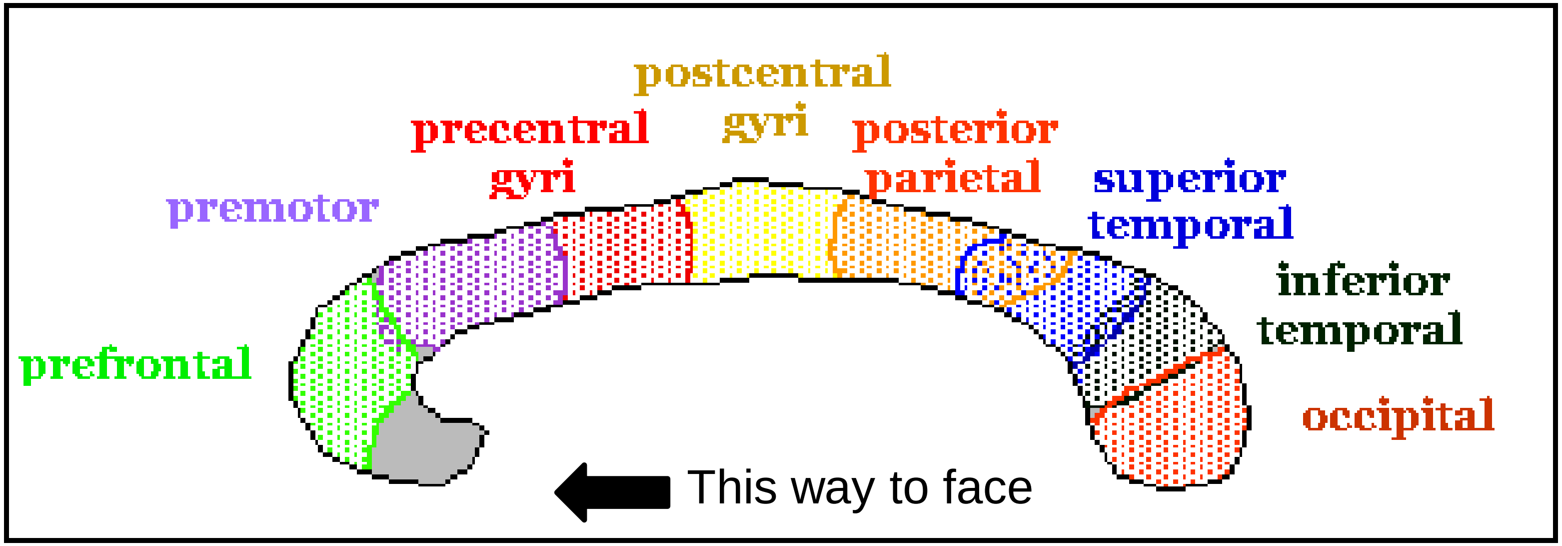
Fidelity of outline reconstruction $\frac{\sum_{k=1}^{K}\sum_{j=1}^{J}|Original_{kj} - Reconstructed_{kj}|}{K}$

glm(outline difference ~ method) Smaller-is-better	(Intercept)	Polygon	rFourier	tFourier
Things which shouldn't change				
Scale	18.24 [15.18,21.3]	51.58 [47.25,55.91]	52.99 [48.66,57.32]	119.72 [115.39,124.05]
Orientation	0.04 [0.04,0.04]	0.14 [0.13,0.14]	0.09 [0.08,0.09]	0.2 [0.2,0.2]
Jitter in outline	0.04 [0.04,0.04]	0.1 [0.1,0.11]	0.1 [0.1,0.11]	0.16 [0.16,0.16]
Randomness in outline	0.04 [0.04,0.04]	0.1 [0.09,0.1]	0.09 [0.09,0.1]	0.19 [0.19,0.2]
Number of points	0.04 [0.04,0.04]	0.1 [0.1,0.1]	0.07 [0.07,0.08]	0.17 [0.17,0.17]
Things which should change				
Midbody length	0.05 [0.04,0.05]	0.1 [0.1,0.11]	0.08 [0.08,0.08]	0.17 [0.17,0.17]
Midbody curve	0.04 [0.04,0.04]	0.11 [0.1,0.11]	0.07 [0.07,0.07]	0.09 [0.09,0.09]
Midbody regularity	0.04 [0.04,0.04]	0.1 [0.1,0.1]	0.09 [0.09,0.09]	0.16 [0.16,0.16]
Splenium size	0.05 [0.05,0.05]	0.1 [0.1,0.1]	0.08 [0.08,0.08]	0.15 [0.15,0.15]
Multiple parameters	0.04 [0.04,0.04]	0.1 [0.1,0.1]	0.08 [0.08,0.08]	0.12 [0.12,0.12]
Real data				
Real Shapes	1.69 [1.61,1.77]	3.31 [3.19,3.43]	2.39 [2.28,2.51]	3.37 [3.25,3.48]

Base group: eFourier. Harmonics: 5, grid: 5x5. n=1000, except for real shapes (n=168). [denotes 95% CI.

Corpus callosum connectivity

- Caveat: thicker doesn't always mean more connections (it also depends on how densely packed it is)
- **Approximate** locations of connections to neural regions



Davis, K. L., Libon, D. J., Nissanov, J., Skalina, S. M., Lamar, M., & Chute, D. L. (1999). Neuropsychological assessment and volumetric magnetic resonance imaging of the corpus callosum in dementia. *Archives of Clinical Neuropsychology*, *14*(8), 622-623.