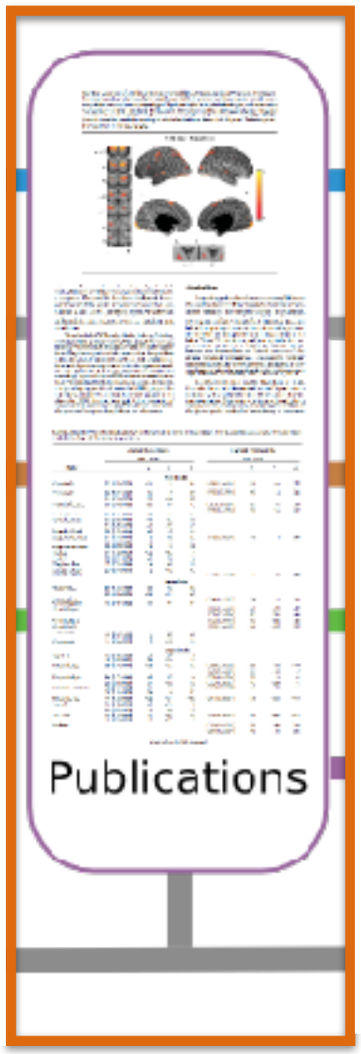
# The dataflow in brain imaging




*Poline, Breeze, Ghosh, et al. (2012)*

# Example 1: Knowledge piece(s)

Brain Lang. 2013 Jun 29. pii: S0093-934X(13)00111-9. doi: 10.1016/j.bandl.2013.05.013. [Epub ahead of print]

## Disrupted white matter in language and motor tracts in developmental stuttering.

Connally EL[1], Ward D, Howell P, Watkins KE.

**Author information**

**Abstract**

White matter tracts connecting areas involved in speech and motor control were examined using diffusion-tensor imaging in a sample of people who stutter (n=29) who were heterogeneous with respect to age, sex, handedness and stuttering severity. The goals were to replicate previous findings in developmental stuttering and to extend our knowledge by evaluating the relationship between white matter differences in people who stutter and factors such as age, sex, handedness and stuttering severity. We replicated previous findings that showed reduced integrity in white matter underlying ventral premotor cortex, cerebral peduncles and posterior corpus callosum in people who stutter relative to controls. Tractography analysis additionally revealed significantly reduced white matter integrity in the arcuate fasciculus bilaterally and the left corticospinal tract and significantly reduced connectivity within the left corticobulbar tract in people who stutter. Region-of-interest analyses revealed reduced white matter integrity in people who stutter in the three pairs of cerebellar peduncles that carry the afferent and efferent fibers of the cerebellum. Within the group of people who stutter, the higher the stuttering severity index, the lower the white matter integrity in the left angular gyrus, but the greater the white matter connectivity in the left corticobulbar tract. Also, in people who stutter, handedness and age predicted the integrity of the corticospinal tract and peduncles, respectively. Further studies are needed to determine which of these white matter differences relate to the neural basis of stuttering and which reflect experience-dependent plasticity.

# Example 1: Variation in white matter in stuttering

Variation in locations of reported white matter differences between people who stutter and people with fluent speech

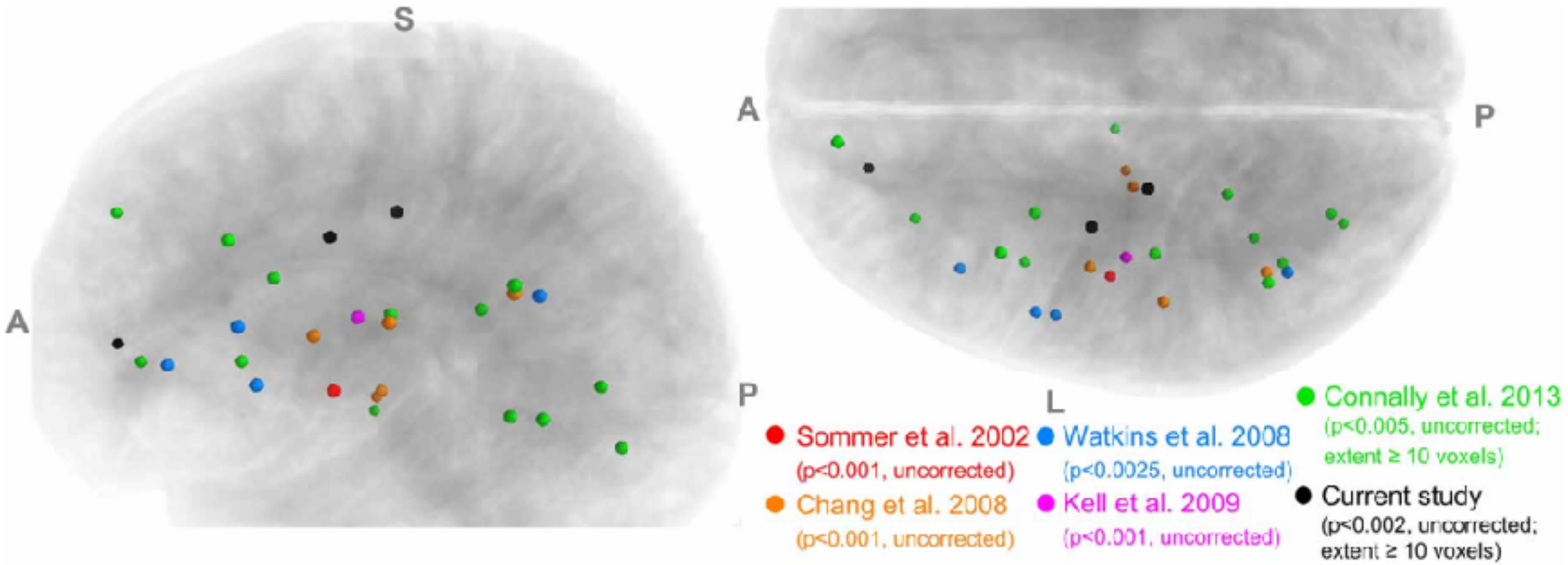



**FIGURE 1 | A summary of voxels with significantly lower FA in PWS than in normal controls reported in five previous studies and the current study.** Only the results in the left hemisphere are shown. **Left panel**: left view; **Right panel**: superior view.

*Cai, Tourville, Perkell, Guenther, Ghosh (2014)*

# Example 2: Knowledge piece(s)

J Neurosci. 2006 Mar 1;26(9):2424-33.

## Cortical and subcortical contributions to Stop signal response inhibition: role of the subthalamic nucleus.

Aron AR[1], Poldrack RA.

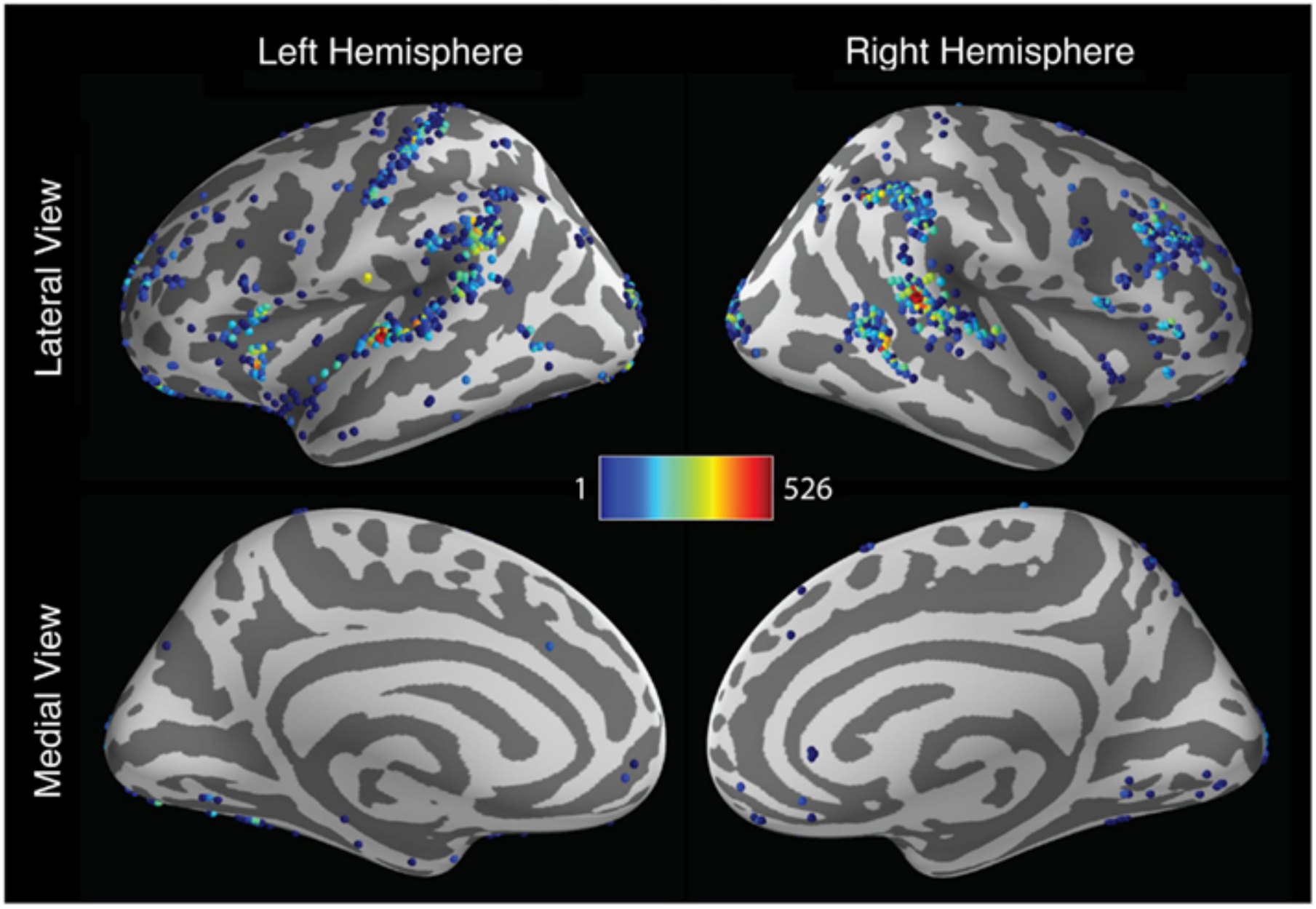
**Author information**

**Abstract**

Suppressing an already initiated manual response depends critically on the right inferior frontal cortex (IFC), yet it is unclear how this inhibitory function is implemented in the motor system. It has been suggested that the subthalamic nucleus (STN), which is a part of the basal ganglia, may play a role because it is well placed to suppress the "direct" fronto-striatal pathway that is activated by response initiation. In two experiments, we investigated this hypothesis with functional magnetic resonance imaging and a Stop-signal task. Subjects responded to Go signals and attempted to inhibit the initiated response to occasional Stop signals. In experiment 1, Going significantly activated frontal, striatal, pallidal, and motor cortical regions, consistent with the direct pathway, whereas Stopping significantly activated right IFC and STN. In addition, Stopping-related activation was significantly greater for fast inhibitors than slow ones in both IFC and STN, and activity in these regions was correlated across subjects. In experiment 2, high-resolution functional and structural imaging confirmed the location of Stopping activation within the vicinity of the STN. We propose that the role of the STN is to suppress thalamocortical output, thereby blocking Go response execution. These results provide convergent data for a role for the STN in Stop-signal response inhibition. They also suggest that the speed of Go and Stop processes could relate to the relative activation of different neural pathways. Future research is required to establish whether Stop-signal inhibition could be implemented via a direct functional neuroanatomic projection between IFC and STN (a "hyperdirect" pathway).

# Example 2: Variation in functional analysis

Variation in functional activity when the same task data was analyzed by different workflows




*Carp (2013)*

# Example 3: Knowledge piece(s)

J Autism Dev Disord. 2012 Nov;42(11):2312-22. doi: 10.1007/s10803-012-1478-z.

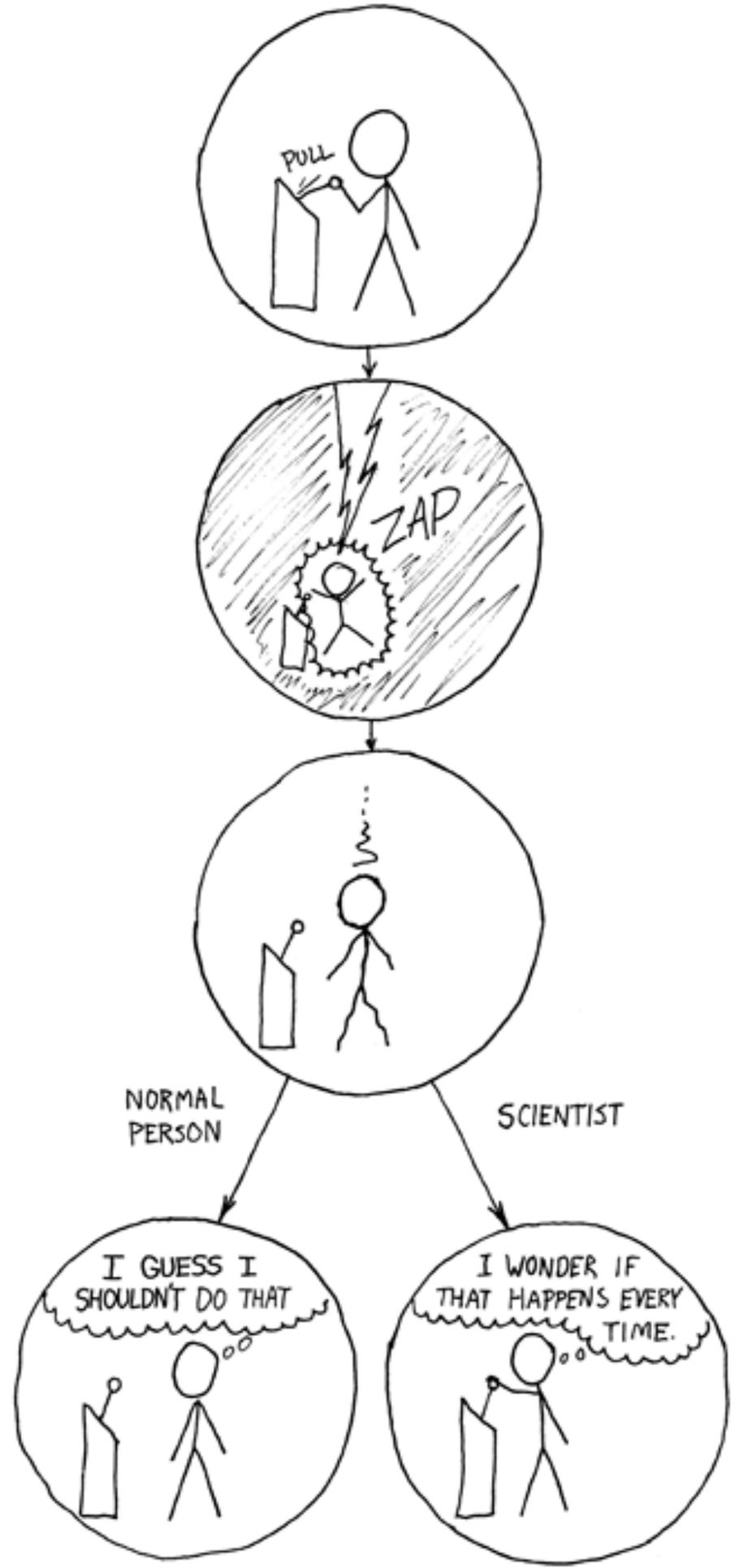
**A two-year longitudinal MRI study of the corpus callosum in autism.**

Frazier TW[1], Keshavan MS, Minshew NJ, Hardan AY.

**Author information**

**Abstract**

A growing body of literature has identified size reductions of the corpus callosum (CC) in autism. However, to our knowledge, no published studies have reported on the growth of CC volumes in youth with autism. Volumes of the total CC and its sub-divisions were obtained from 23 male children with autism and 23 age- and gender-matched controls at baseline and 2-year follow-up. Persistent reductions in total CC volume were observed in participants with autism relative to controls. Only the rostral body subdivision showed a normalization of size over time. Persistent reductions are consistent with the diagnostic stability and life-long impairment observed in many individuals with autism. Multi-modal imaging studies are needed to identify specific fiber tracks contributing to CC reductions.


PULL
ZAP
NORMAL PERSON
SCIENTIST
I GUESS I SHOULDN'T DO THAT
I WONDER IF THAT HAPPENS EVERY TIME.

xkcd:242

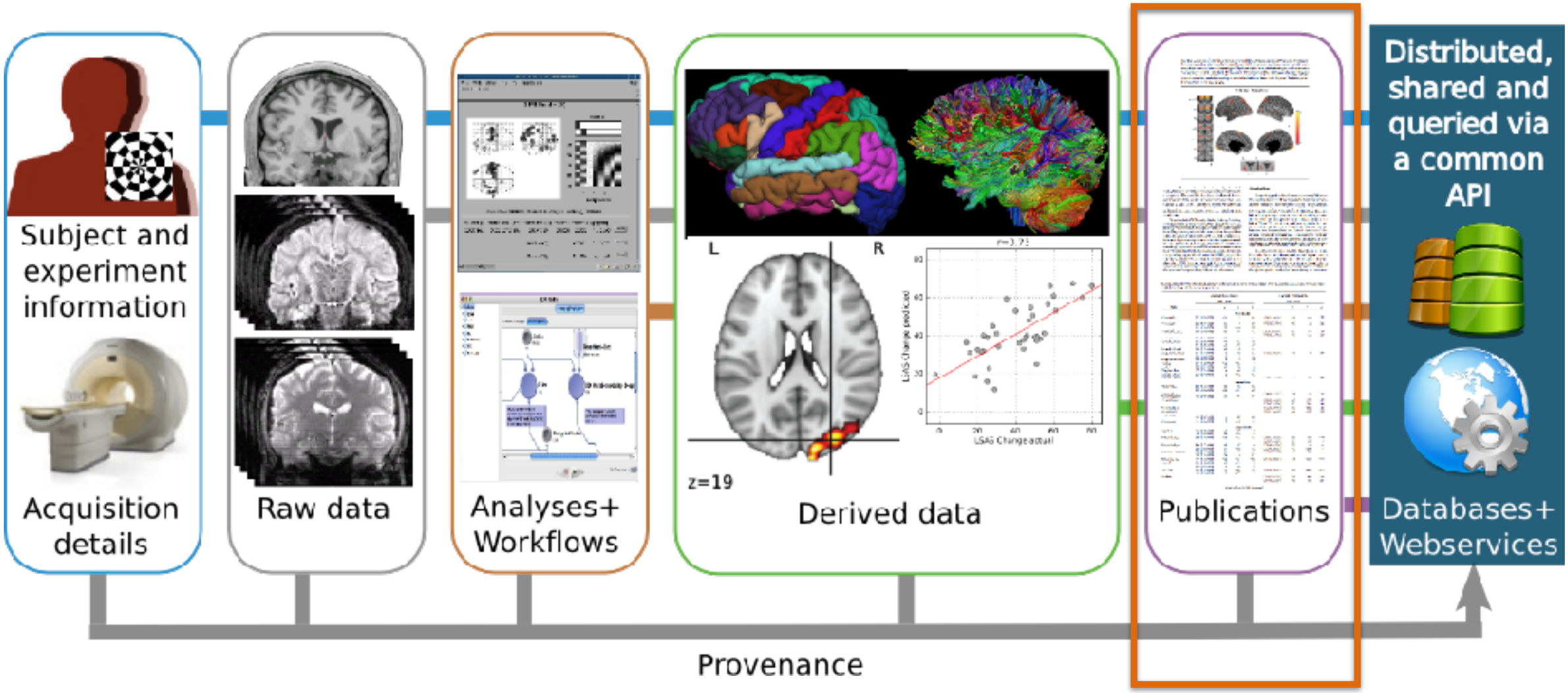
# Example 3: Underpowered studies

Corpus callosum volume difference between ASD and controls could not be replicated in a cohort of 694 participants.

| | Study | $Mean_{ASD} \pm SD$ ($cm^2$) | $Mean_{CTRL} \pm SD$ ($cm^2$) | Pooled SD ($cm^2$) | Effect Size | P-Value (2-sided) | Achieved power (2-sided) | Power to detect SD=0.3 (2-sided) |
|---|---|---|---|---|---|---|---|---|
| 1 | Gaffney 1987 (34) | 5.89±1.04 | 6.24±1.37 | 1.31 | -0.27 | 0.070 | 14.6% | 17.3% |
| 2 | Egaas 1995 (35) | 5.57±0.99 | 5.89±0.91 | 0.95 | -0.33 | **0.0011** | 38.2% | 32.3% |
| 3 | Piven 1996 (36) | 6.15±0.83 | 6.40±0.38 | 0.64 | -0.39 | **0.0017** | 36.9% | 24.1% |
| 4 | Elia1999 (37) | 5.26±1.00 | 5.41±0.64 | 0.91 | -0.16 | 0.36 | 6.4% | 12.7% |
| 5 | Manes 1999 (38) | 4.64±0.99 | 5.71±0.97 | 0.99 | -1.07 | **<0.0001** | 93.6% | 16.2% |
| 6 | Rice 2005 (39) | 7.34±1.11 | 7.75±1.14 | 1.15 | -0.35 | 0.13 | 11.4% | 9.3% |
| 7 | Vidal 2006 (40) | 6.06±1.15 | 6.68±0.79 | 0.99 | -0.62 | **<0.0001** | 57.7% | 17.9% |
| 8 | Boger 2006 (41) | 4.59±0.67 | 4.99±0.72 | 0.70 | -0.57 | **<0.0001** | 67.6% | 24.1% |
| 9 | Alexander 2007 (42) | 7.87±1.99 | 9.32±1.70 | 1.88 | -0.77 | **<0.0001** | 91.4% | 25.2% |
| 10 | Just 2007 (3) | 6.40±0.88 | 7.1±0.88 | 0.89 | -0.78 | **<0.0001** | 62.8% | 13.9% |
| 11 | Hardan 2009 (43) | 5.74±1.13 | 6.58±1.04 | 1.10 | -0.76 | **<0.0001** | 69.8% | 16.2% |
| 12 | Freitag 2009 (44) | 6.22±0.45 | 6.54±1.24 | 0.96 | -0.34 | 0.071 | 14.6% | 12.2% |
| 13 | Keary 2009 (45) | 6.19±1.09 | 6.76±1.10 | 1.11 | -0.51 | **<0.0001** | 54.0% | 22.4% |
| 14 | Anderson 2011 (46) | 6.54±1.20 | 7.05±0.90 | 1.09 | -0.46 | **<0.0001** | 59.4% | 29.6% |
| 15 | Hong 2011 (47) | 8.14±1.31 | 8.27±1.27 | 1.31 | -0.10 | 0.58 | 4.6% | 13.3% |
| 16 | Frazier 2012 (48) | 6.30±1.11 | 6.78±1.08 | 1.11 | -0.43 | **0.005** | 29.9% | 16.7% |
| 17 | Prigge 2013 (49) | 5.74±0.91 | 6.24±0.89 | 0.91 | -0.55 | **<0.0001** | 83.7% | 36.0% |

*Lefebvre, Beggiato, Bourgeron, Toro (2014)*

# The dataflow in brain imaging




*Poline, Breeze, Ghosh, et al. (2012)*

# Ubiquity of error in scientific computing

"In my own experience, error is ubiquitous in scientific computing, and one needs to work very diligently and energetically to eliminate it. One needs a very clear idea of what has been done in order to know where to look for likely sources of error.

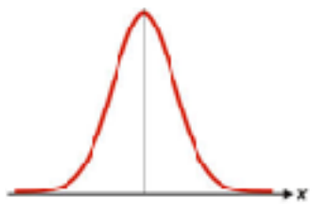
I often cannot really be sure what a student or colleague has done from his/her own presentation, and in fact often his/her description does not agree with my own understanding of what has been done, once I look carefully at the scripts.

Actually, I find that researchers quite generally forget what they have done and misrepresent their computations."

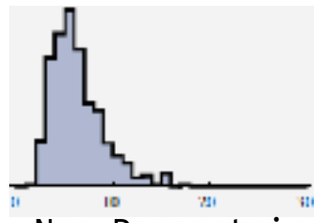
*Donoho (2010)*

# Standardization across software

- Model of the error
  - Prob. distribution:

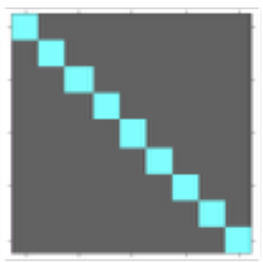



Gaussian

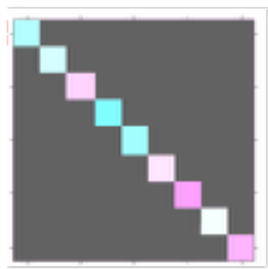


Non-Parametric

  - Variance:

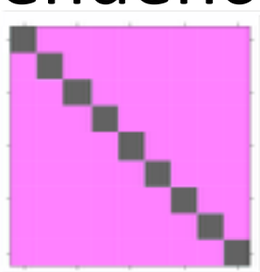


homogeneous

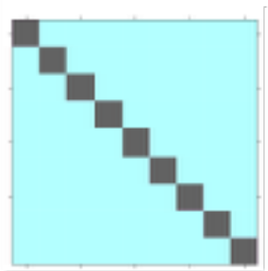


heterogeneous

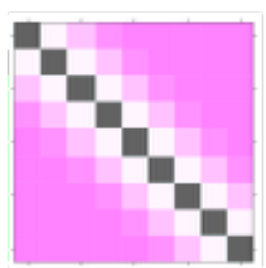
  - Dependence:



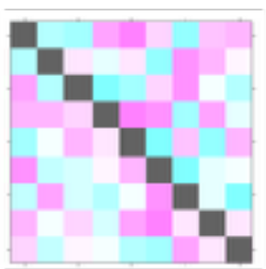
Independent noise



Compound Symmetry



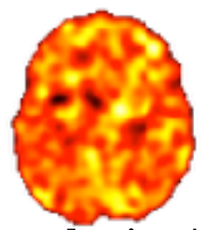
Serially correlated



Arbitrarily correlated

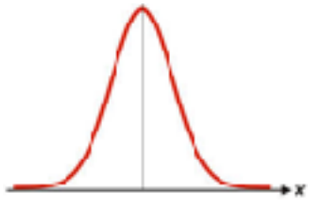
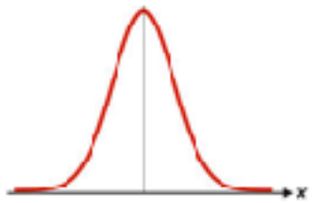
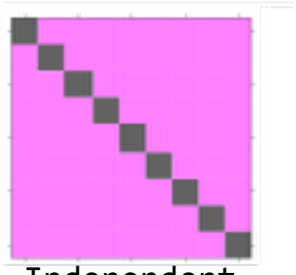
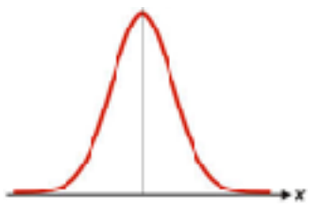
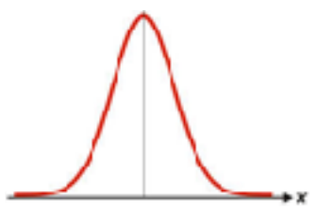
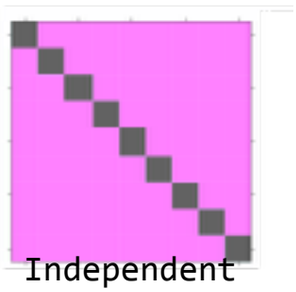
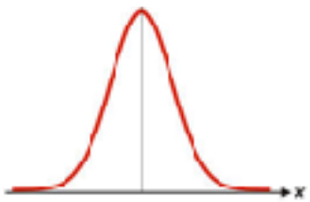
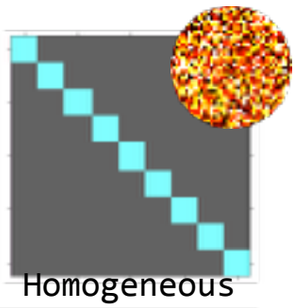
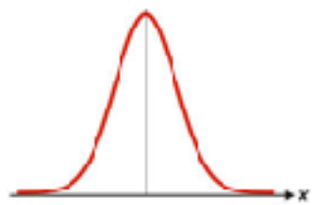
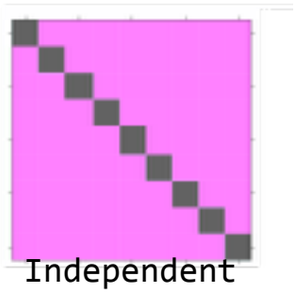
global

local

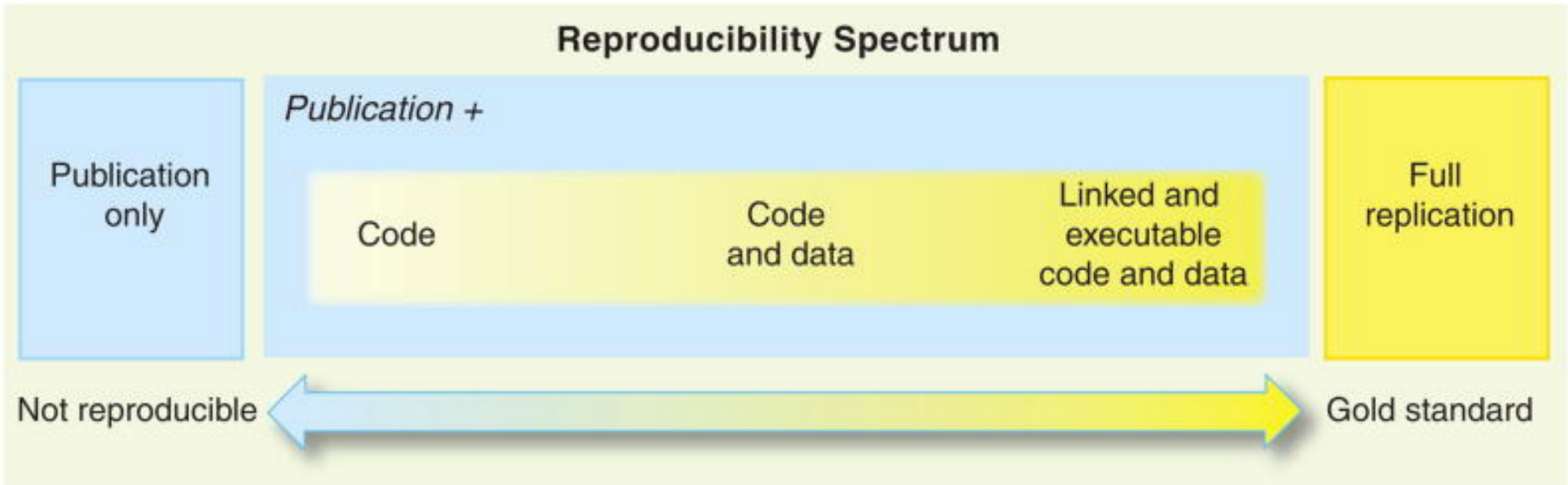


regularized

# Error Models

| | 1st level | | | 2nd level | | |
|---|---|---|---|---|---|---|
| SPM |  Gaussian |  Homogeneous local |  Serial. corr. global |  Gaussian |  Homogeneous local |  Independent noise |
| FSL |  Gaussian |  Homogeneous local |  Serial. corr. regularized |  Gaussian |  Heterogeneous local |  Independent noise |
| AFNI |  Gaussian |  Homogeneous local |  Serial. corr. local |  Gaussian |  Hetero- or Homogeneous local |  Independent noise |

# The Reproducibility Matrix




**Repeatable** Can the same lab repeat all steps of an experiment?

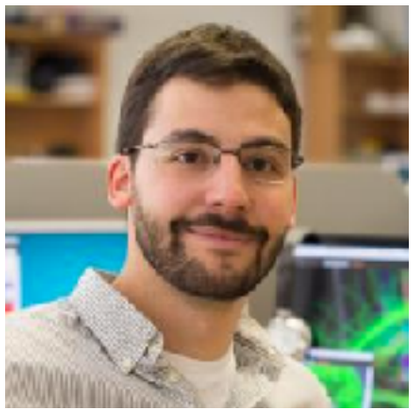
**Replicable** Can another lab redo the experiment?

**Reproducible** Can we get similar results - input to output relation?

**Reusable** Can someone reuse your data/code/samples/hardware?

*Peng (2011), Science; Goble (2014), GigaScience; Drummond (2009), ICML;*

djarecka



oesteban

mgxd

chrisfilo

effigies

Nipype is currently supported by 1R01 EB020740

Nipype: Dataflows for Reproducible Biomedical Research.