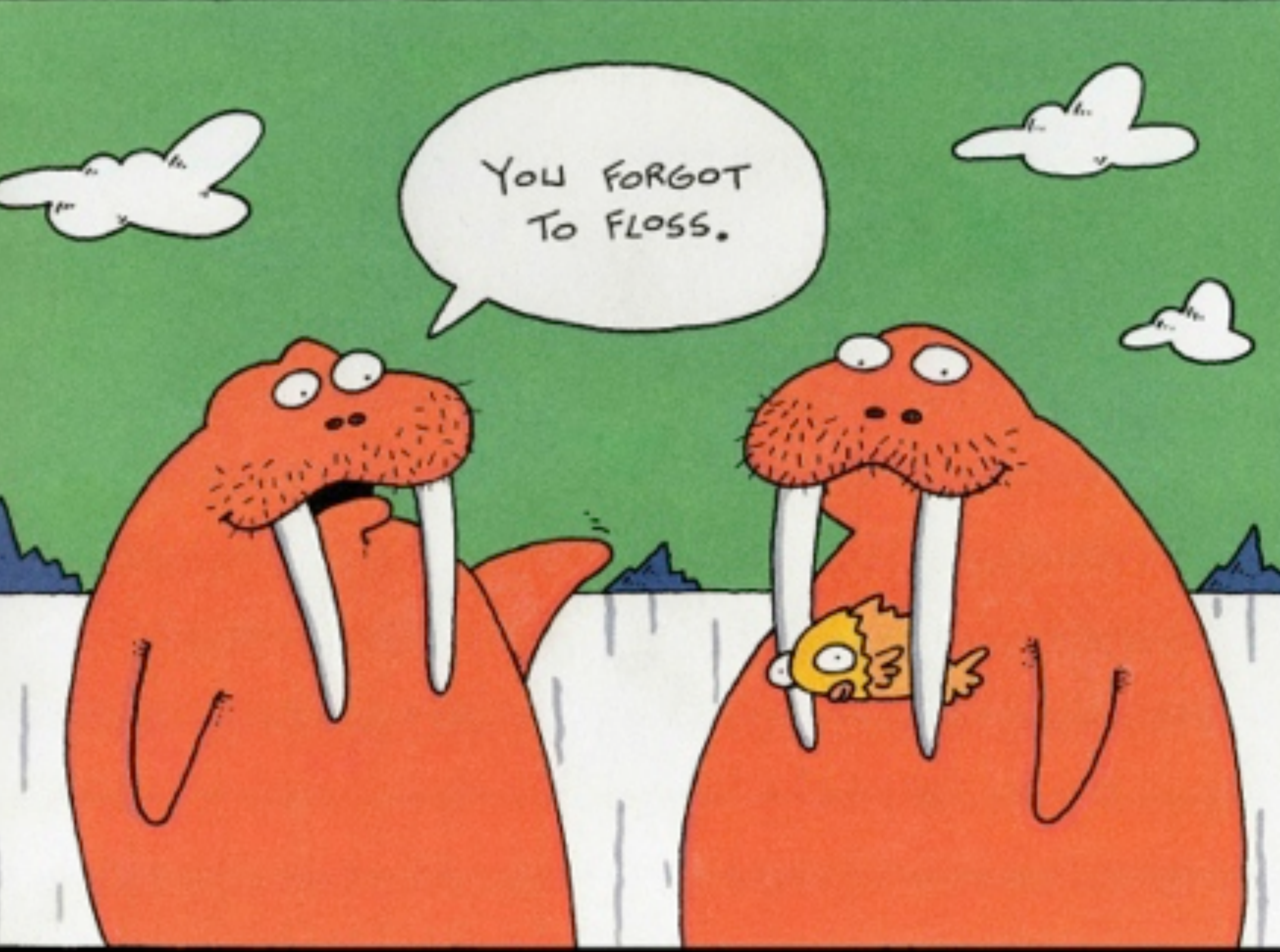
NeuroVault and the vision for data sharing in neuroimaging

Chris Gorgolewski
Max Planck Research Group: Neuroanatomy & Connectivity
Leipzig, Germany

Data sharing?


YOU FORGOT TO FLOSS.

Data sharing?

- Ok, ok so we should share data.
- We all know it's good.
- But almost no one does it.
  - You have to prepare data
  - You risk that your mistakes will be found!

Part I: Large scale data sharing is a fact

Part II: Data sharing does not have to be expensive

Part III: Making data sharing count

Part IV: Implications of data sharing

The data out there is calling you…

PART I: LARGE SCALE DATA SHARING IS A FACT

NKI Enhanced

- 329 subjects (will reach 1000)
  - Representative sample: young and old, some with mental health history
- 1 hour worth of MRI (3T) scanning:
  - MPRAGE (TR = 1900; voxel size = 1mm isotropic)
  - 3x resting state scans (645msec, 1400msec, and 2500msec)
  - Diffusion Tensor Imaging (137 direction; voxel size = 2mm isotropic)
  - Visual Checkboard and Breath Holding manipulations

General Information

- Demographic Questionnaire (DEMOS)*
- Edinburgh Handedness Inventory (EHI)
- Hollingshead Four Factor Index of Socioeconomic Status (SES)*
- Medical History Questionnaire (Med-Hist)
- Medical Conditions Questionnaire*
- Medications Questionnaire*

Physical Measures

- Actigraphy
- Bike Test*
- Blood draw: chemistry profile, lipid profile, thyroid profile, CBC with differential, lead level, genetics, pregnancy test*
- Urine sample (Drug Test)(11+)*
- Height/Weight*
- Hip/Wait Measurements*
- Ishihara Color Vision Test (Color)*
- MRI Mock Scan*
- MRI Scan*
- Tanner Staging (TANN)(6-17)
- Vital Signs*

Neurocognitive Tasks*

- Attention Network Task (ANT)
- Computerized Neurocognitive Battery (CNB)
- Delis-Kaplan Executive Function System (D-KEFS)(8+)
- Grip strength
- Grooved Purdue Pegboard
- Wechsler Abbreviated Scale of Intelligence-II (WASI-II)
- Wechsler Individual Achievement Test-II-Abbreviated (WIAT-II-A)

Diagnostic Assessments*

- Adult ADHD Clinical Diagnostic Scale (ACDS)(18+)
- Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS-PL)(6-17)
- Structured Clinical Interview for DSM-IV – Non-Patient edition (SCID-NP)(18+)

Behavioral Measures

- 21-Item Peters et al. Delusions Inventory (PDI-21)(13+)
- Autism Spectrum Screening Questionnaire (ASSQ)(6-17)
- Behavioral Assessment System for Children (BASC-2)(6-17)
- Cambridge-Hopkins Restless Leg Syndrome Questionnaire (CHRLS)(13+)
- Child Behavioral Checklist (CBCL)(6-17); Achenbach Youth Self Report (YSR)(11-17); Adult Self Report (ASR)(18-59); Older Adult Self Report (OASR)(60+)
- Child Eating Behavior Questionnaire (CEBQ)(6-11); Three Factor Eating Questionnaire (TFEQ)(12+)
- Children's Behavior Questionnaire (CBQ)(6-8); Early Adolescent Temperament Questionnaire Parent Report (EATQ)(9-15); Adult Temperament Questionnaire (ATQ)(16+)
- Children's Depression Inventory (CDI-II)(7-17); Beck Depression Inventory-II (BDI)(18-64); Geriatric Depression Scale-Long Form (GDS-LF)(65+)
- Children's Yale-Brown Obsessive Compulsive Scale (CY-BOCS)(6-17)*; Yale-Brown Obsessive Compulsive Scale (Y-BOCS)(18+)*
- Cognitive Failures Questionnaire (CFQ)(15+)
- Comprehensive Adolescent Severity Inventory – Alcohol and Other Drugs (CASI-AOD)(11+)
- Conners' Parent Rating Scale-Revised-Short (CPRS-R-S)(6-17); Conner-Wells' Adolescent Self-Report Scale- Short (CASS-S)(8-17); Conners Adult ADHD Rating Scales (CAARS)(18+)
- DOSPERT Risk Taking Scale (DOSPERT)(18+)
- Eating Disorder Examination Questionnaire (EDEQ)(13+)
- Fagerstrom Tolerance Questionnaire for Adolescents (FTAQ)(13-17); Fagerstrom Test for Nicotine Dependence (FTND)(18+)
- International Physical Activity Questionnaire (IPAQ)(15+)
- Interpersonal Reactivity Index (IRI)(13+)
- Inventory of Callous-Unemotional Traits Parent Report (ICU-P)(6-17); Inventory of Callous-Unemotional Traits Youth Version (ICU-Y)(13+)
- Multidimensional Anxiety Scale for Children (MASC)(8-17); State Trait Anxiety Inventory (STAI)(18+)
- NEO Five Factor Inventory (NEO-FFI)(12+)
- Pittsburgh Sleep Quality Index (PSQI)(13+)
- Repetitive Behavior Scale-Revised (RBS-R)(6-17)
- Social Responsiveness Scale, Parent Report (SRS)(6-17)
- Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behavior Scale Parent Version (SWAN)(6-17)
- Trauma Symptom Checklist for Children (TSC-C)(8-17); Trauma Symptom Checklist (TSC-40)(18+)
- UCLA PTSD Reaction Index – Parent version (UCLA-RI-P)(6-17); UCLA PTSD Reaction Index for Children and Adolescents (UCLA-RI)(8+)
- UPPS Impulsive Behavior Scale (UPPS-P)(18+)
- Vineland Adaptive Behavior Scales Parent Rating Form, Second Edition (Vineland-II)(6+)*
- Yale Global Tic Severity Scale (YGTSS)(6+)*
- Youth Risk Behavior Surveillance System (YRBSS)(11-21)

Assessments for all participants ages 6 years and older unless indicated
(*) terms: Clinician administered

fcon_1000.projects.nitrc.org/indi/enhanced/

Human Connectome Project

- ~232 subjects (will reach 1200)
  - Young and healthy (22-35yrs)
  - 200 twins!
- 1 hour worth of MRI scanning:
  - Resting-state fMRI (R-fMRI)
  - Task-evoked fMRI (T-fMRI)
    - Working Memory
    - Gambling
    - Motor
    - Language
    - Social Cognition
    - Relational Processing
    - Emotion Processing
  - Diffusion MRI (dMRI)

Human Connectome Project

- Rich phenotypical data
  - Cognition, personality, substance abuse etc.
- Genotyping! (not yet available)
- Methodological developments
  - Fine tuned sequences
  - New preprocessing techniques
- Ready to use preprocessed data

humanconnectome.org

Test-retest datasets

- NKI multiband Test-retest
  fcon_1000.projects.nitrc.org/indi/pro/eNKI_RS_TRT/FrontPage.html
- Classification learning and stop-signal (1 year test-retest)
  openfmri.org/dataset/ds000017
- A test-retest fMRI dataset for motor, language and spatial attention functions
  www.gigasciencejournal.com/content/2/1/6

FCP/INDI Usage Survey

FCP/INDI Data Usage Description

Master's thesis research	11.94%
Doctoral dissertation research	38.81%
Teaching resource (projects or examples)	13.43%
Pilot data for grant applications	16.42%
Research intended for publication	76.12%
Independent study (e.g., teach self about analysis)	37.31%

FCP/INDI Users; 10% respondent rate

Survey Courtesy of Stan Colcombe & Cameron Craddock

Sharing little things…

PART II: DATA SHARING DOES NOT HAVE TO BE EXPENSIVE

Just coordinates?

- Databases such as Neurosynth or BrainMap rely on peak coordinates reported in papers (only strong effects)

Spatial memory task	MNI coordinates (mm)			
	x	*y*	*z*	*Z* **stat**
Z > 2.0				
Subcortical regions				
Right thalamus	18	–14	8	2.60
Right pallidum	22	–4	2	2.98
Right putamen	30	–20	0	3.51
Left thalamus	–12	–14	10	3.44
Left pallidum	–18	–4	–2	3.34
Left caudate	–12	4	10	3.06

Are we throwing money away?

Baby steps

- Everything is a question of cost and benefit
  - If we keep the cost low even small benefit (or just conviction that data sharing is GOOD) will suffice

NeuroVault.org
simple data sharing

- Minimize the cost!
- We just want your statistical maps with minimum description (DOI)
  - If you want you can put more metadata, but you don't have to
- We streamline login process (Google, Facebook)

Benefits?

- In return authors get interactive web based visualization of their statistical maps
  - Something they can embed on their lab website
- We are keeping both cost and benefit low…
  - …but we also plan to work with journal editors to popularize the idea

Live demo

Using NeuroVault…

- Improves collaboration
- Makes your paper more attractive
- Shows you care about transparency
- Takes only five minutes
- Gives you warm and fuzzy feeling that you helped future meta-analyses

NeuroVault for developers

- RESTful API (field tested by Neurosynth)
- Source code available on GitHub
  www.github.com/chrisfilo/NeuroVault

NeuroVault.org

Credit where credit's due

PART III: MAKING DATASHARING COUNT

Motivation

vs.

institutions

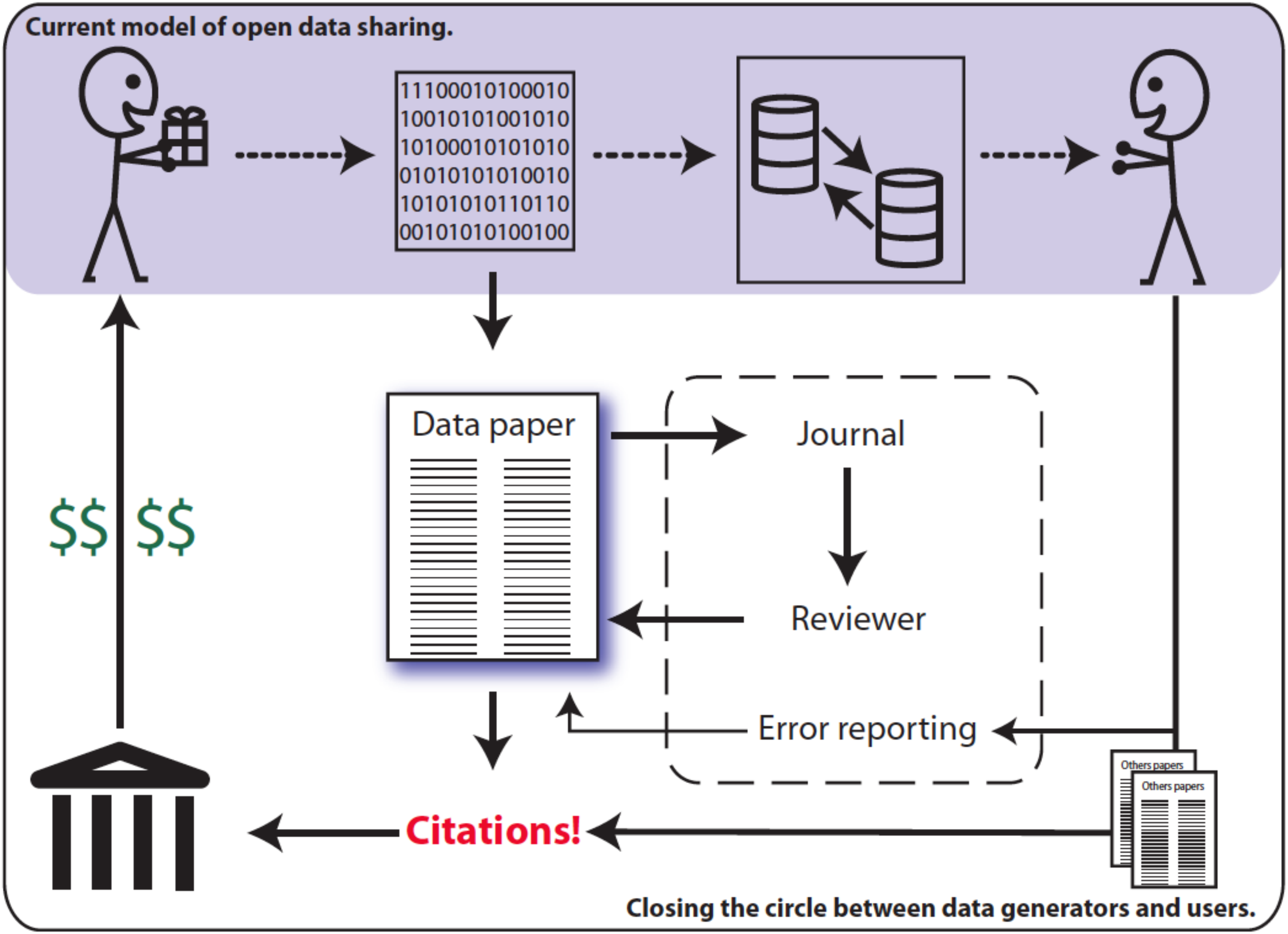
scientists

Quality control

Complex datasets require elaborate descriptions

Solution – data papers

- Authors get recognizable credit for their work.
  - Even smaller contributors such as RAs can be included.
- Acquisition methods are described in detail.
- Quality of metadata is being controlled by peer review.



Gorgolewski, Milham, and Margulies, 2013

Where to publish data papers?

- Neuroinformatics (Springer)
- Frontiers in Human Brain Methods (Frontiers Media)
- GigaScience (BGI, BioMed Central)
- Scientific Data (Nature Publising Group, coming soon)

Where to publish data papers?

- Neuroinformatics (Springer)
- Frontiers in Human Brain Methods (Nature Publishing Group)
- GigaScience (BGI, BioMed Central)
- Scientific Data (Nature Publising Group, coming soon)

Gorgolewski *et al. GigaScience* 2013, **2**:6
http://www.gigasciencejournal.com/content/2/1/6

DATA NOTE
Open Access

A test-retest fMRI dataset for motor, language and spatial attention functions

Krzysztof J Gorgolewski[1*], Amos Storkey[1], Mark E Bastin[2], Ian R Whittle[3], Joanna M Wardlaw[2] and Cyril R Pernet[2]

Abstract

Background: Since its inception over twenty years ago, functional magnetic resonance imaging (fMRI) has been used in numerous studies probing neural underpinnings of human cognition. However, the between session variance of many tasks used in fMRI remains understudied. Such information is especially important in context of clinical applications. A test-retest dataset was acquired to validate fMRI tasks used in pre-surgical planning. In particular, five task-related fMRI time series (finger, foot and lip movement, overt verb generation, covert verb generation, overt word repetition, and landmark tasks) were used to investigate which protocols gave reliable single-subject results. Ten healthy participants in their fifties were scanned twice using an identical protocol 2–3 days apart. In addition to the fMRI sessions, high-angular resolution diffusion tensor MRI (DTI), and high-resolution 3D T1-weighted volume scans were acquired.

Findings: Reliability analyses of fMRI data showed that the motor and language tasks were reliable at the subject level while the landmark task was not, despite all paradigms showing expected activations at the group level. In addition, differences in reliability were found to be mostly related to the tasks themselves while task-by-motion interaction was the major confounding factor.

Conclusions: Together, this dataset provides a unique opportunity to investigate the reliability of different fMRI tasks, as well as methods and algorithms used to analyze, de-noise and combine fMRI, DTI and structural T1-weighted volume data.

Keywords: Test-retest, Overt verb generation, Covert verb generation, Overt word repetition, Landmark, Motor, fMRI, DTI www.nature.com/scientificdata

SCIENTIFIC DATA

SUBJECT CATEGORIES
» Induced pluripotent stem cells
» Proteomic analysis
» Microarray analysis

Proteomic profiles of human embryonic stem cells, induced-pluripotent stem cells and precursor fibroblasts

Javier Munoz[1] and Albert J.R. Heck[2,3]

Assessing relevant molecular differences between human-induced pluripotent stem cells (hiPSCs) and human embryonic stem cells (hESCs) is important, given that such differences may impact their potential therapeutic use. Controversy surrounds recent gene expression studies comparing hiPSCs and hESCs. Here, we present a dataset comprising quantitative mass spectrometry-based measurements of the proteomes of hESCs, two different hiPSCs and their precursor fibroblast cell lines, along with matching gene expression profiles for each sample. These data are suitable for in depth comparative analysis of the proteomes of both somatic and pluripotent cells, and have been deposited in three different public repositories to maximize ease of reuse by the community.

Design Type(s)	cell type comparison design • growth condition intervention design
Measurement Type(s)	protein expression profiling • transcription profiling assay
Technology Type(s)	mass spectrometry assay • DNA microarray
Factor Type(s)	cell line • growth condition
Sample Characteristic(s)	*Homo sapiens* • embryonic stem cell line • embryonic fibroblast cell line • foreskin fibroblast cell line

PART IV: IMPLICATIONS OF DATASHARING

Sample sizes will grow

By combining multiple shared datasets or using one of the big datasharing initiatives we will gain access to bigger sample

Bigger samples

- Better parameter estimates
- Lower ratio of false positives (and false negatives)
- Lower risk of inflated effect sizes
- Higher power: better sensitivity to small effects

Is more power bad?

- In classical hypothesis testing the null hypothesis usually states no difference
- However nothing in nature is exactly the same
- In most cases we just don't have enough power to see it
- Some differences are more important than others

Sex differences in the structural connectome of the human brain

Madhura Ingalhalikar[a,1], Alex Smith[a,1], Drew Parker[a], Theodore D. Satterthwaite[b], Mark A. Elliott[c], Kosha Ruparel[b], Hakon Hakonarson[d], Raquel E. Gur[b], Ruben C. Gur[b], and Ragini Verma[a,2]

[a]Section of Biomedical Image Analysis and [c]Center for Magnetic Resonance and Optical Imaging, Department of Radiology, and [b]Department of Neuropsychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; and [d]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104

Edited by Charles Gross, Princeton University, Princeton, NJ, and approved November 1, 2013 (received for review September 9, 2013)

Sex differences in human behavior show adaptive complementarity: Males have better motor and spatial abilities, whereas females have superior memory and social cognition skills. Studies also show sex differences in human brains but do not explain this complementarity. In this work, we modeled the structural connectome using diffusion tensor imaging in a sample of 949 youths (aged 8–22 y, 428 males and 521 females) and discovered unique sex differences in brain connectivity during the course of development. Connection-wise statistical analysis, as well as analysis of regional and global network measures, presented a comprehensive description of network characteristics. In all supratentorial regions, males had greater within-hemispheric connectivity, as well as enhanced modularity and transitivity, whereas between-hemispheric connectivity and cross-module participation predominated in females. However, this effect was reversed in the cerebellar connections. Analysis of these changes developmentally demonstrated differences in trajectory between males and females mainly in adolescence and in adulthood. Overall, the results suggest that male brains are structured to facilitate connectivity between perception and coordinated action, whereas female brains are designed to facilitate communication between analytical and intuitive processing modes.

(12–14 y) (23), and this result was established on a larger sample size (114 subjects) as well (24). On the other hand, sex differences on the entire age range (childhood to old age) demonstrated higher FA and lower MD in males (19, 25, 26). Similar findings of higher FA in males were obtained with tractography on major WM tracts (27, 28).

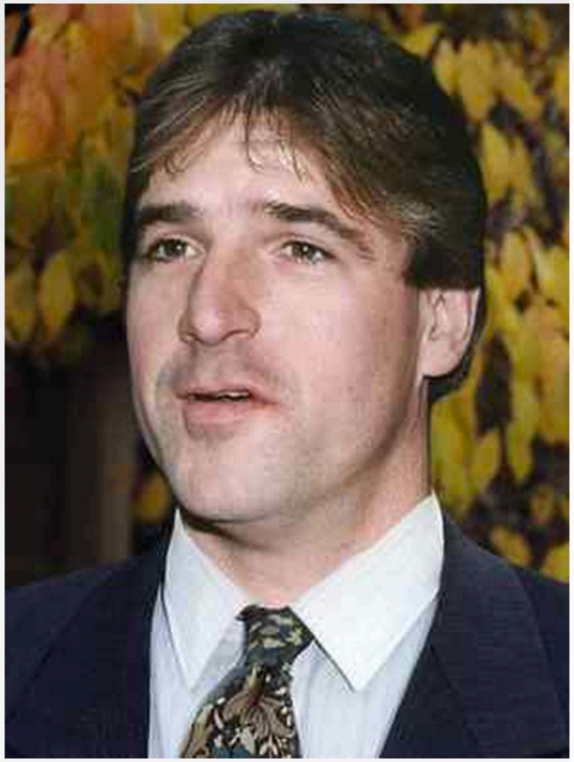
Rather than investigating individual regions or tracts in isolation, the brain can be analyzed on the whole as a large and complex network known as the human connectome (29). This connectome has the capability to provide fundamental insights into the organization and integration of brain networks (30). Advances in fiber tractography with diffusion imaging can be used to understand complex interactions among brain regions and to compute a structural connectome (SC) (31). Similar functional connectomes (FCs) can be computed using modalities like functional MRI, magnetoencephalography, and EEG. Differences in FCs have revealed sex differences and sex-by-hemispheric interactions (32), with higher local functional connectivity in females than in males (33). Although SCs of genders have displayed small-world architecture with broad-scale characteristics (34, 35), sex differences in network efficiency have been reported (36), with women having greater overall cortical con-

Male and female brains wired differently, scans reveal

Maps of neural circuitry show women's brains are suited to social skills and memory, men's perception and co-ordination

Gender differences hard wired

The hardwired difference between male and female brains could explain why men are 'better at map reading'



"I'm sure this study can't possibly be misinterpreted in any way."

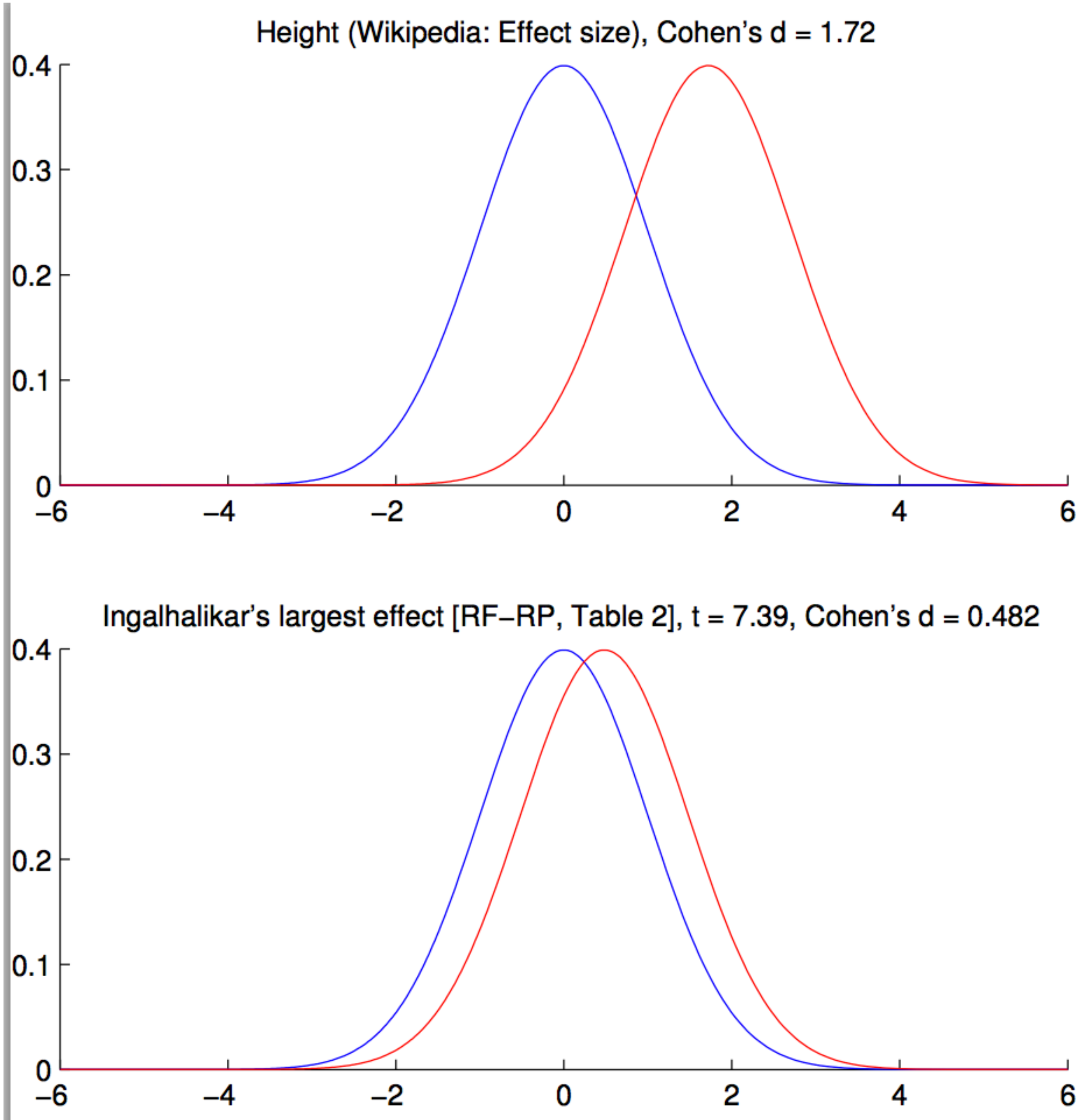
Jackson Haney –
Adhesive Sprayer

"I usually just go by what kind of genitalia they have."

Anna Duarte –
Bookkeeper

"So who wins?"

Mike Lucero –
Instrument Repair Supervisor



Ridgway, Gerard (2013): Illustrative effect sizes for sex differences. figshare.
http://dx.doi.org/10.6084/m9.figshare.866802

Is more power bad?

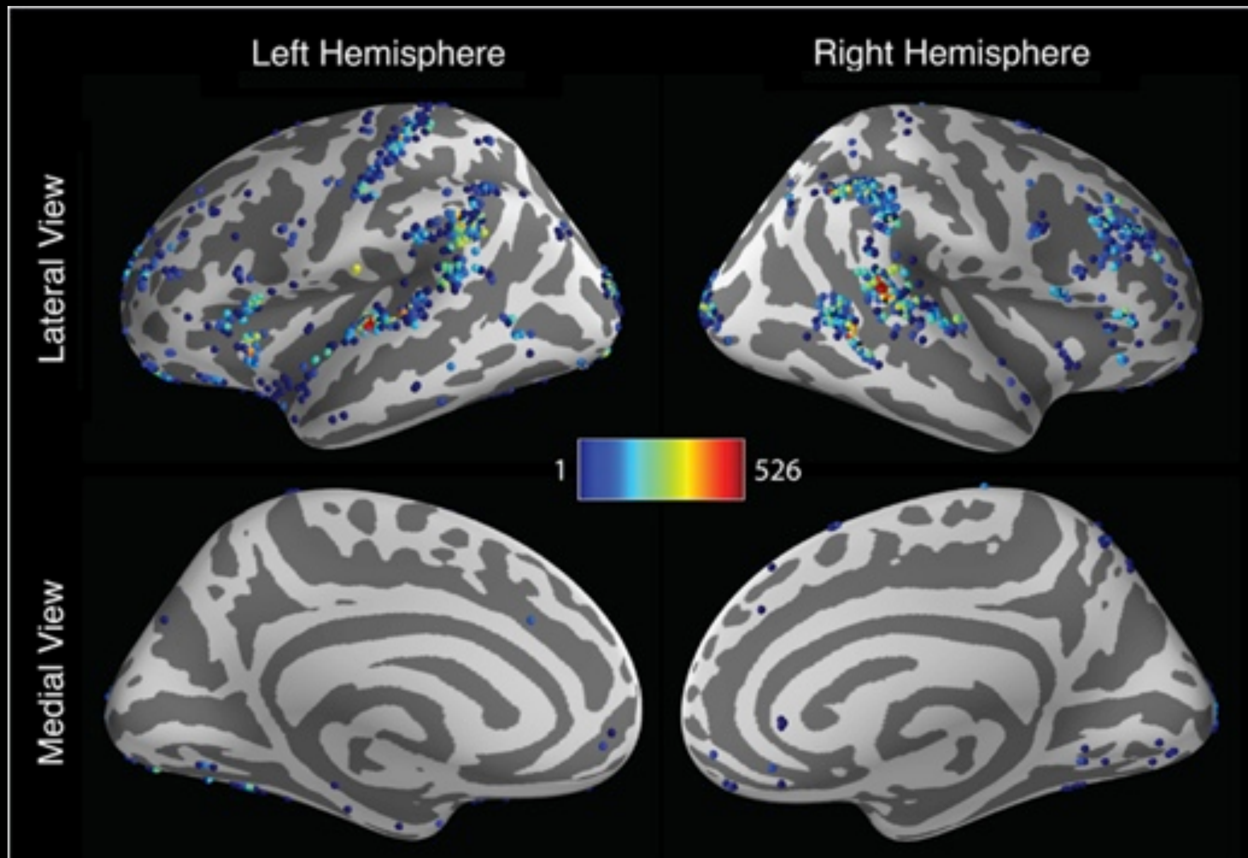
- No – an “overpowered study” is an oxymoron
- But we will need to revise our methods
- Incorporate our assumption of what is a trivial effect size in our analyses
  - Either through Bayesian framework or different null hypotheses
- Start looking at effect and confidence interval maps instead of just thresholded p-maps

Vibration effect

- While analyzing an MRI dataset we face a plethora of choices
- Some alternatives have no clear bad or good options
- The vibration effect is the ratio of effect size of the highest and lowest effects across all processing options

Ioannidis, J. P. a. (2008). Why most discovered true associations are inflated.

Vibration effect



- Carp J (2012) On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments. Front. Neurosci. 6:149. doi: 10.3389/fnins.2012.00149

The Effects of FreeSurfer Version, Workstation Type, and Macintosh Operating System Version on Anatomical Volume and Cortical Thickness Measurements

Ed H. B. M. Gronenschild[1,2]*, Petra Habets[1,2], Heidi I. L. Jacobs[1,2,3], Ron Mengelers[1,2], Nico Rozendaal[1,2], Jim van Os[1,2,4], Machteld Marcelis[1,2]

1 Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University Medical Center, Maastricht, Alzheimer Center Limburg, The Netherlands, **2** European Graduate School of Neuroscience (EURON), Maastricht University, Maastricht, The Netherlands, **3** Cognitive Neurology Section, Institute of Neuroscience and Medicine-3, Research Centre Jülich, Jülich, Germany, **4** King's College London, King's Health Partners, Department of Psychosis Studies Institute of Psychiatry, London, United Kingdom

Abstract

FreeSurfer is a popular software package to measure cortical thickness and volume of neuroanatomical structures. However, little if any is known about measurement reliability across various data processing conditions. Using a set of 30 anatomical T1-weighted 3T MRI scans, we investigated the effects of data processing variables such as FreeSurfer version (v4.3.1, v4.5.0, and v5.0.0), workstation (Macintosh and Hewlett-Packard), and Macintosh operating system version (OSX 10.5 and OSX 10.6). Significant differences were revealed between FreeSurfer version v5.0.0 and the two earlier versions. These differences were on average 8.8±6.6% (range 1.3–64.0%) (volume) and 2.8±1.3% (1.1–7.7%) (cortical thickness). About a factor two smaller differences were detected between Macintosh and Hewlett-Packard workstations and between OSX 10.5 and OSX 10.6. The observed differences are similar in magnitude as effect sizes reported in accuracy evaluations and neurodegenerative studies. The main conclusion is that in the context of an ongoing study, users are discouraged to update to a new major release of either FreeSurfer or operating system or to switch to a different type of workstation without repeating the analysis; results thus give a quantitative support to successive recommendations stated by FreeSurfer developers over the years. Moreover, in view of the large and significant cross-version differences, it is concluded that formal assessment of the accuracy of FreeSurfer is desirable.

Vibration effect

- We will finally see how much (or little) our analyses replicate over different datasets and **methods**

Take home message(s)

I. Take advantage of shared data
II. Share your statistical maps at NeuroVault.org
III. Share your data and publish a data paper
IV. Expect changes in the way we analyze our data

Acknowledgements
(my personal giants)

Pierre-Louie Bazin	Jonathan Smallwood
Haakon Engen	Yannick Schwarz
Satrajit Ghosh	Florence J.M. Ruby
Russell A. Poldrack	Melaina Vinski
Jean-Baptiste Poline	Camille Maumet
Yannick Schwarz	Dan Lurie
Tal Yarkoni	Sebastian Urchs
Michael Milham	Judy A. Kipping
Daniel Margulies	R. Cameron Craddock
Benjamin Baird	MPI CBS Resting state group

THANK YOU!

“I swear I’ve heard it before”

- In the past there were many attempts to propagate data sharing
  - For example fMRI DC:
    - technical issues
    - …and the amount of time it took to prepare data for submission (a week, a very frustrating week)
- fMRI DC was however very ambitious for its time:
  - They wanted to collect raw data and all metadata required to reproduce the analysis

Van Horn & Gazzaniga (2013). Why share data? Lessons learned from the fMRIDC. NeuroImage