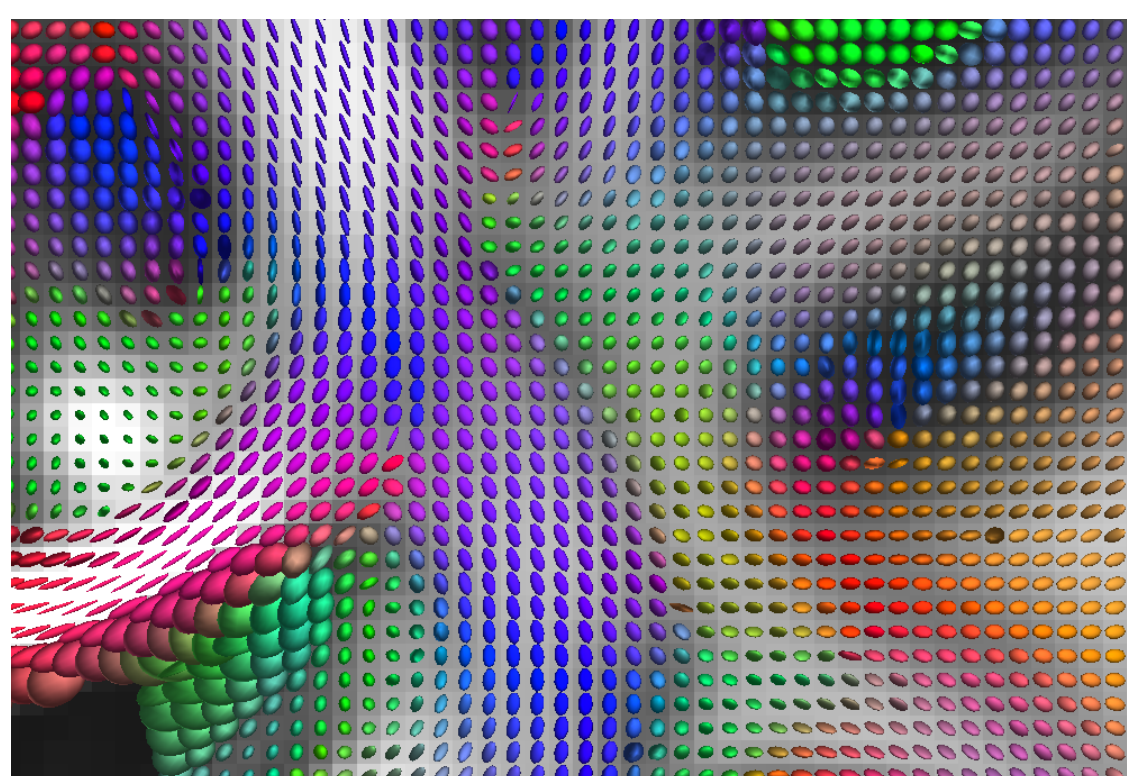
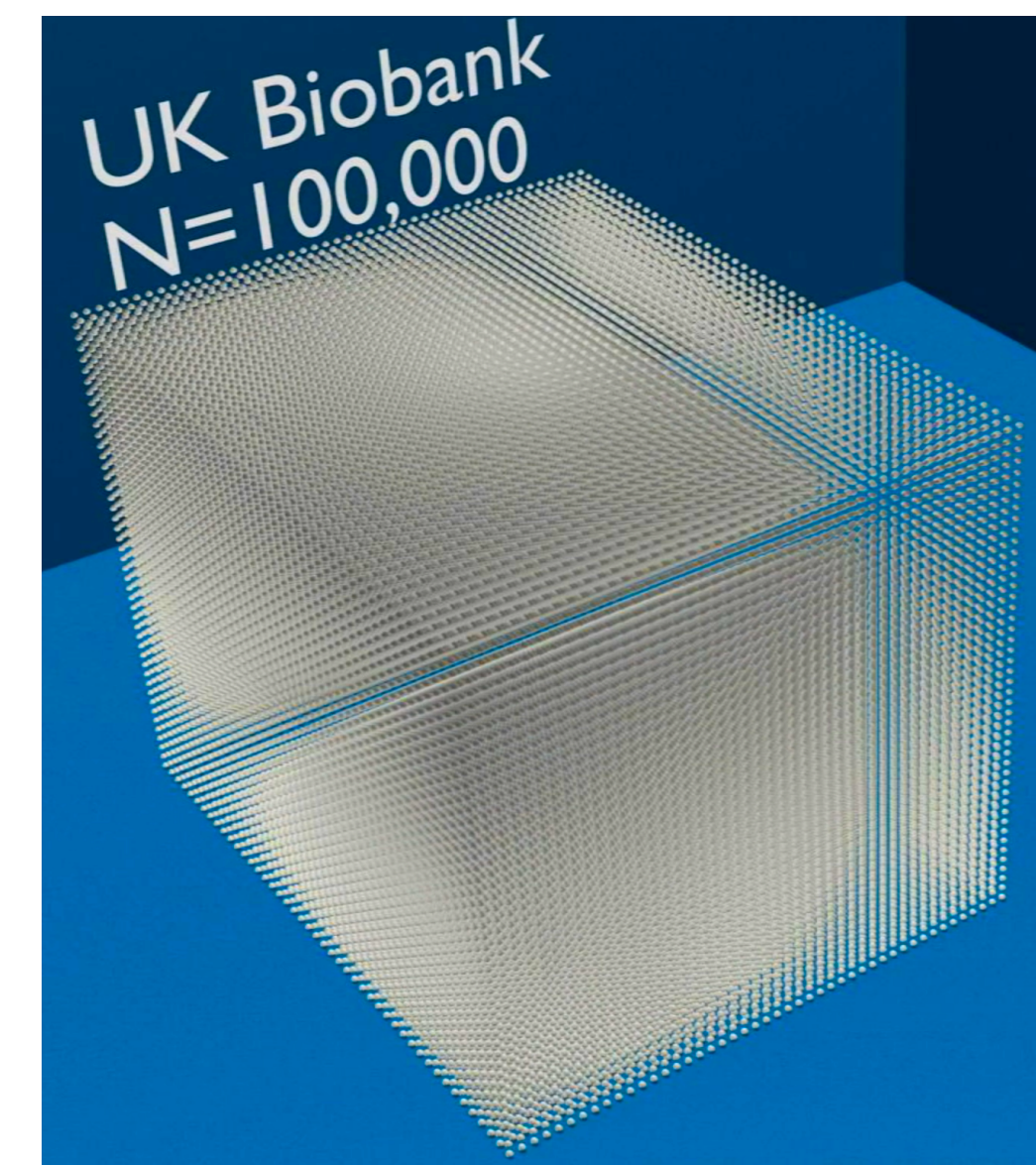
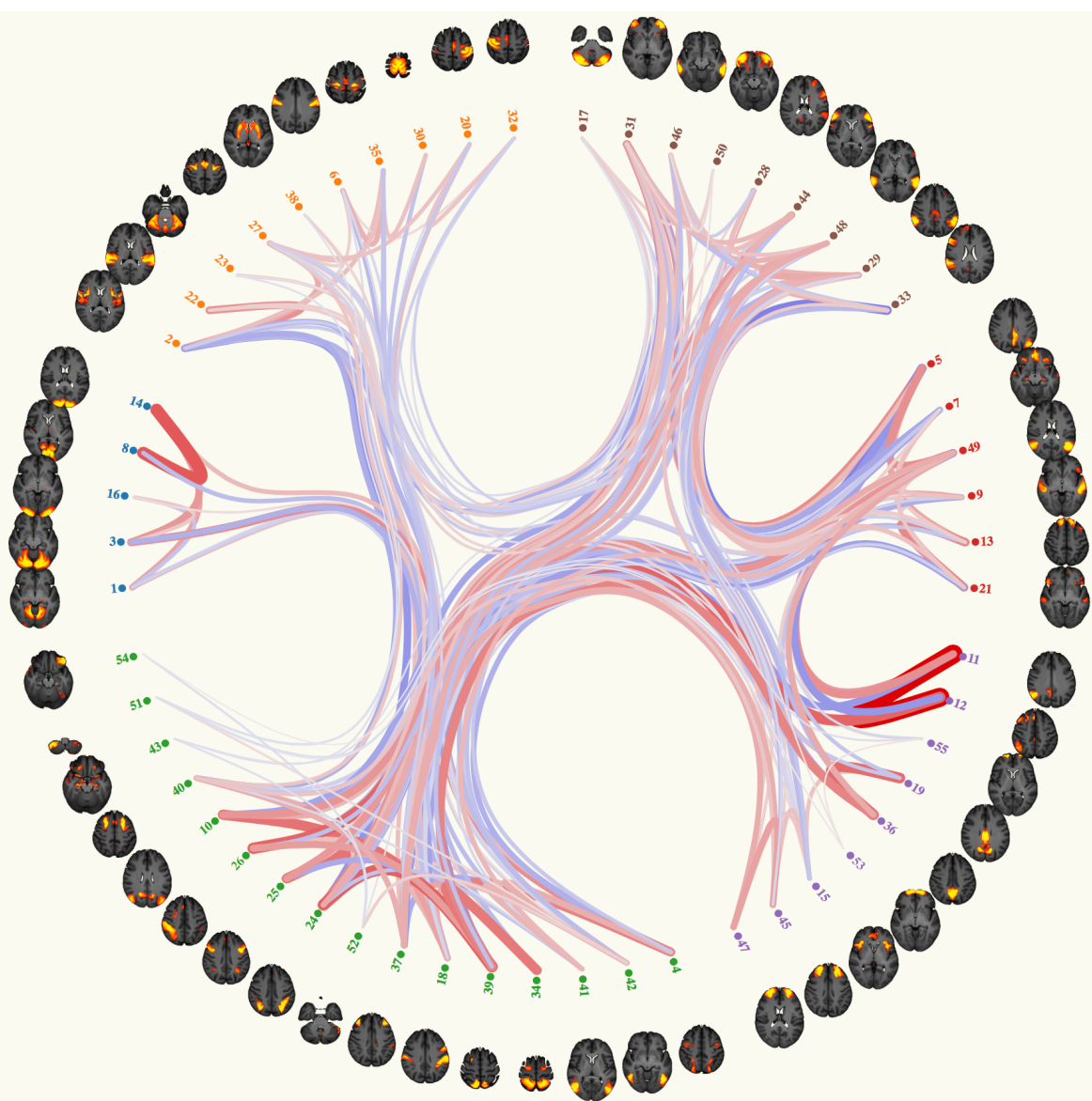


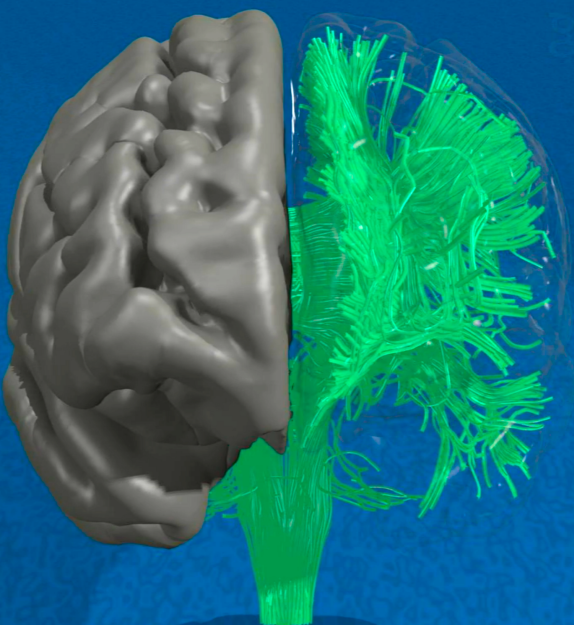
biobank^{uk}
Imaging study

Having fun with FSL in UK Biobank

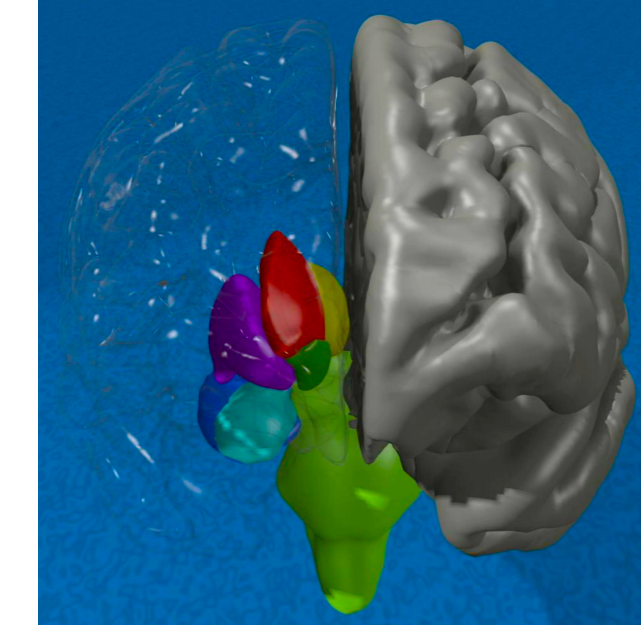
Fidel Alfaro Almagro
WIN, University of Oxford
(a.k.a FMRIB)

[@neurofidelity.bsky.social](https://bsky.app/profile/neurofidelity.bsky.social)





biobank^{uk}



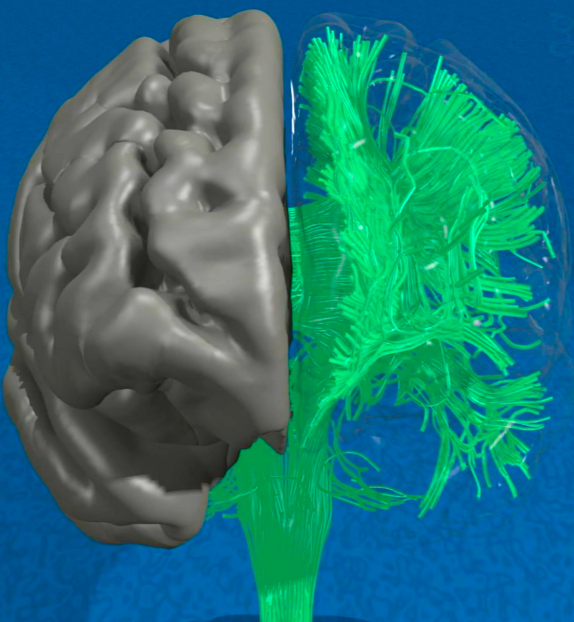
- 1. Background**
2. Brain Imaging Project
3. Brain Imaging: What can we do?



UNIVERSITY OF
OXFORD



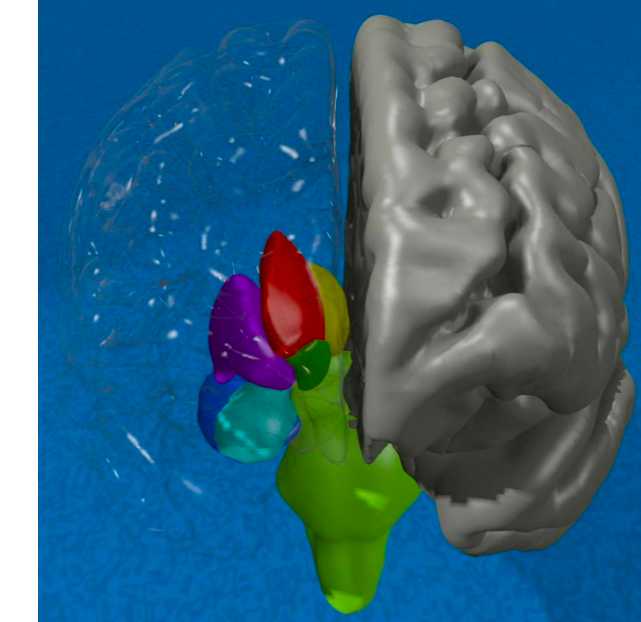
wellcome
centre
integrative
neuroimaging



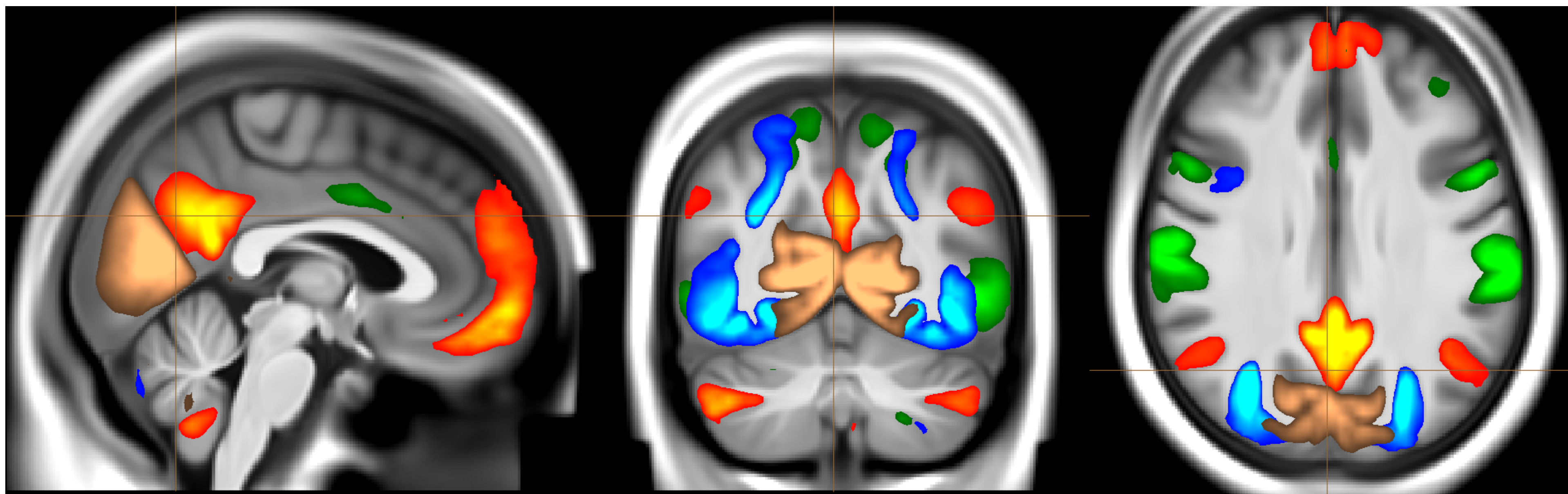
biobank^{uk}

Imaging study

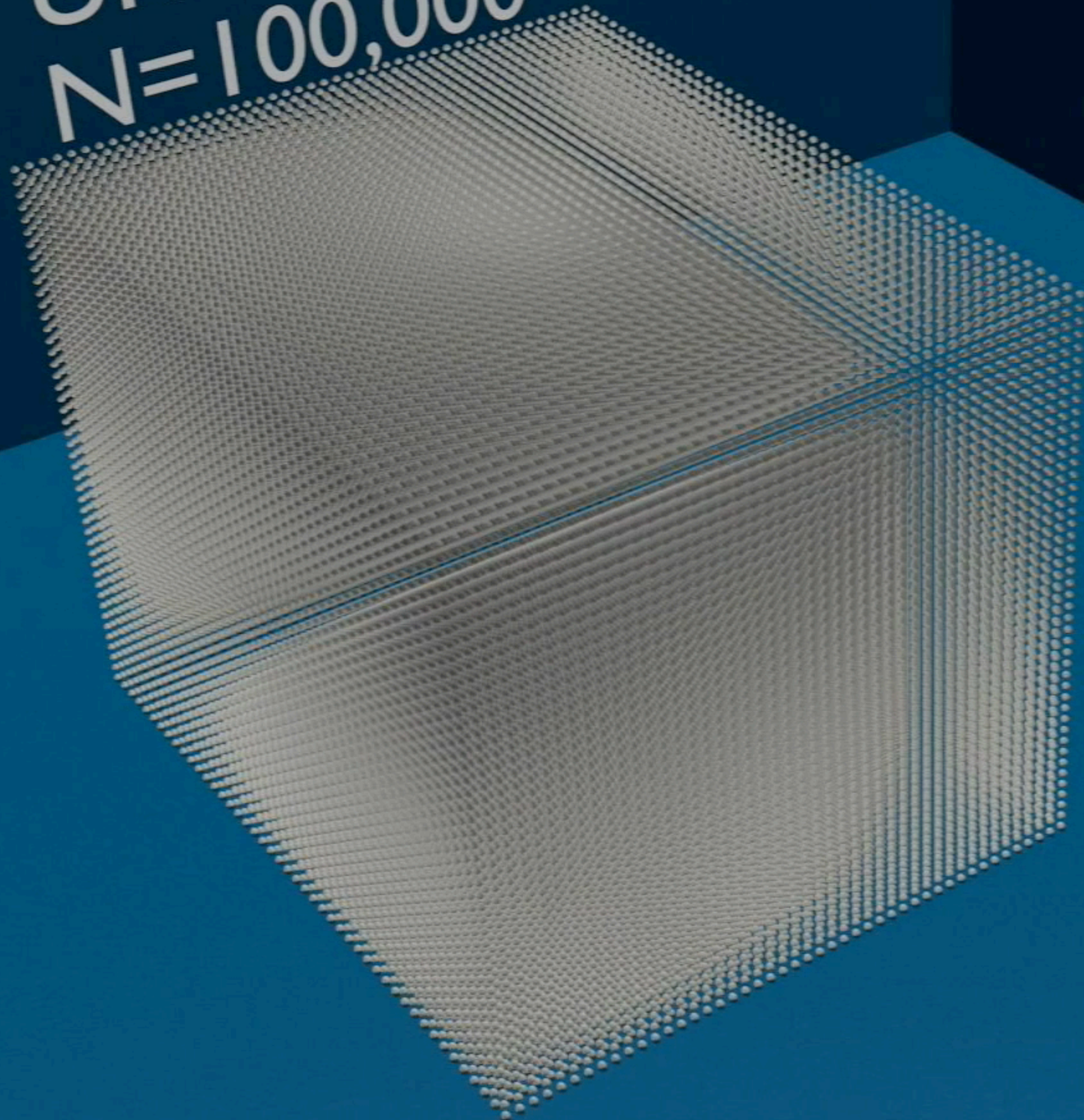
- Brain imaging scientific direction:
Stephen Smith, Karla Miller (Oxford), Paul Matthews (Imperial)
- Brain imaging analysis pipeline:
Fidel Alfaro Almagro, Stephen Smith (Oxford) and many others



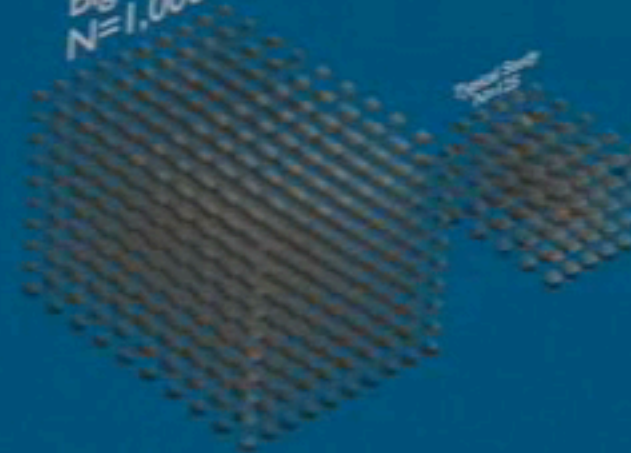
- Prospective epidemiological study: 500,000, 45-75y, UK residents
- Genetic data + biological samples + lifestyle information + health records.
- Discover early markers & risk factors of disease

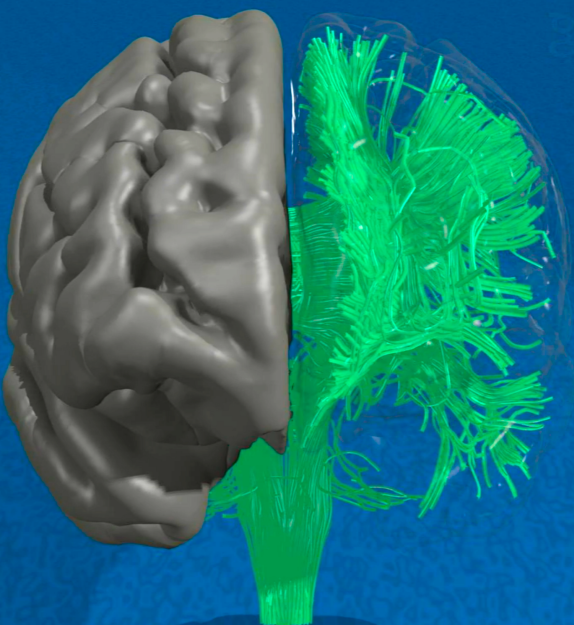


UK Biobank
N=100,000

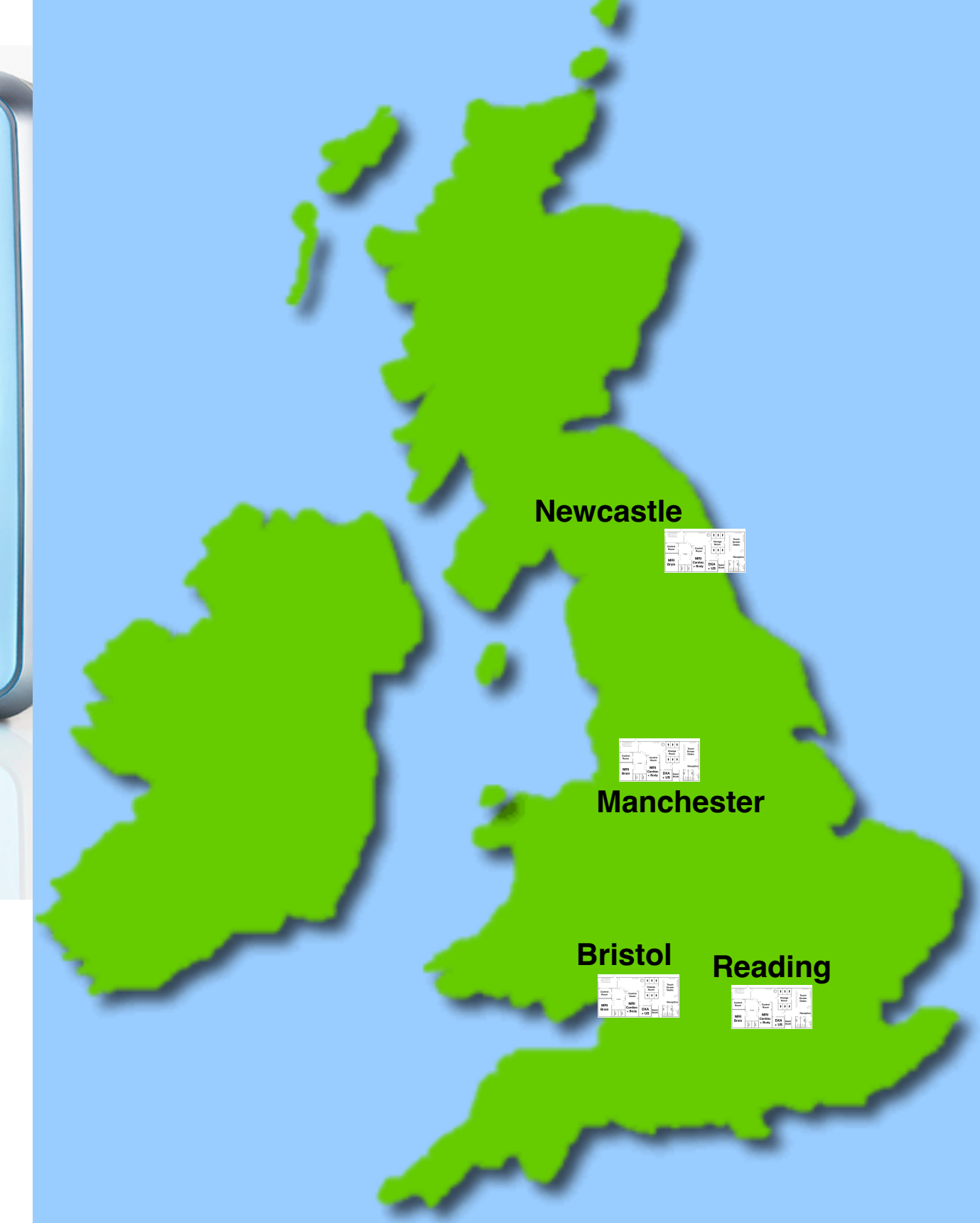


Big Study
N=1,000





biobank^{uk}
Imaging study



Imaging extension

100,000 participants
+ 60,000 2nd timepoint

How do you scan 160,000 subjects?

4 dedicated centres, 72 subjects/day, 7 days/week

What imaging data can
feasibly be acquired
in a 4-hour visit?

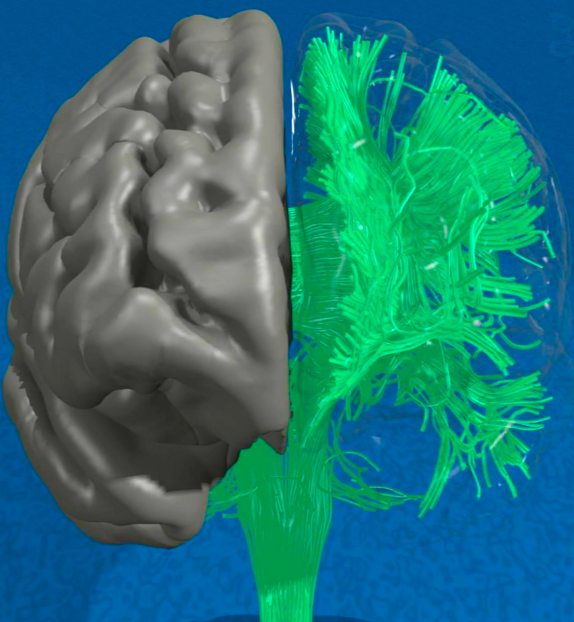
- Brain MRI (3T)
- Cardiac & body MRI (1.5 T)
- DXA bone density
- Carotid UltraSound



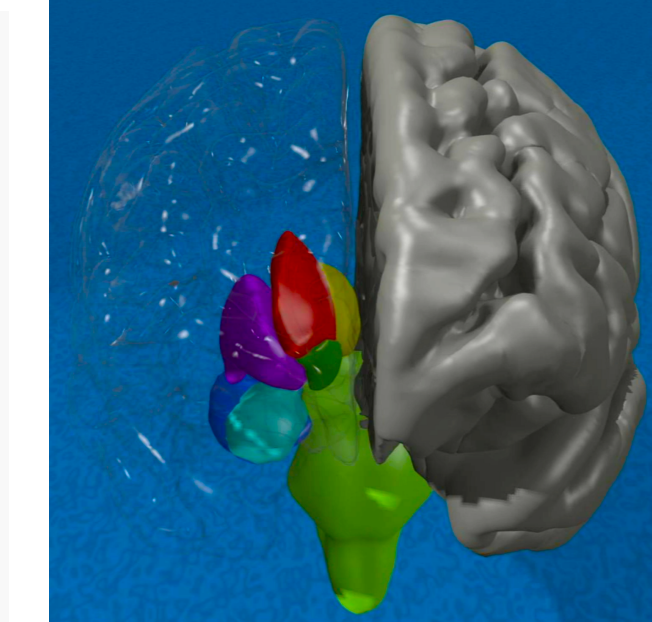
UNIVERSITY OF
OXFORD



wellcome
centre
integrative
neuroimaging

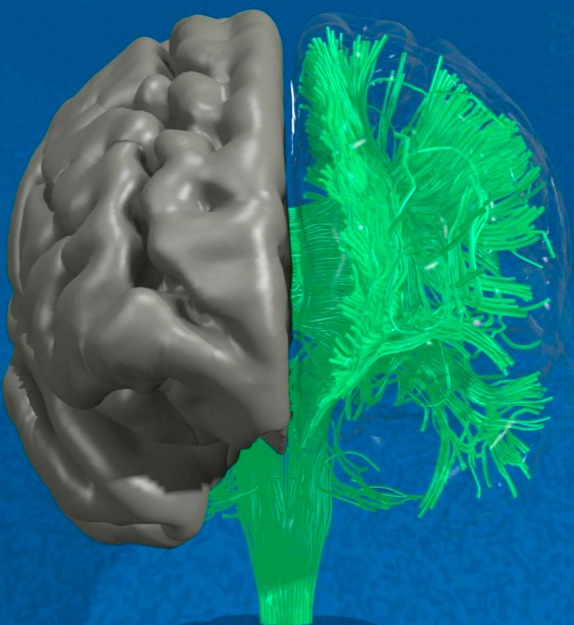


biobank^{uk}
Imaging study

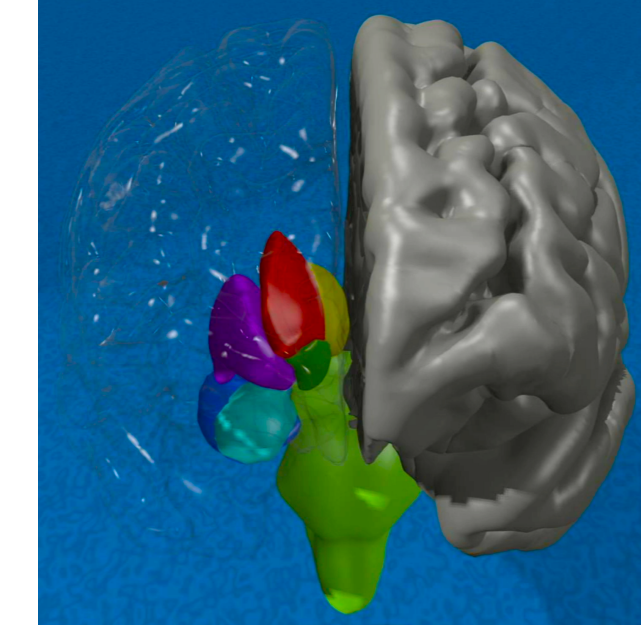


Data Access <http://www.ukbiobank.ac.uk/register-apply>

- Open for use by researchers worldwide
- Application process (protection of sensitive subject data, etc)
- Modest data access fee to fund indefinite maintenance of resource
- No preferential access to scientists helping run UK Biobank !



biobank^{uk}

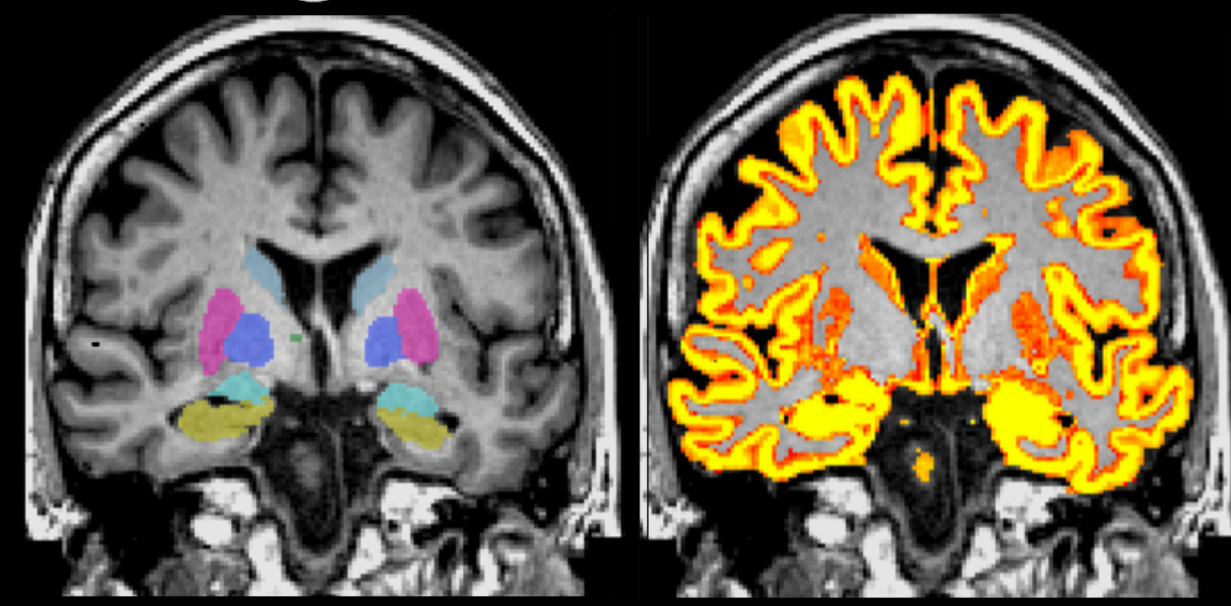


1. Background
- 2. Brain Imaging Project**
3. Brain Imaging: What can we do?

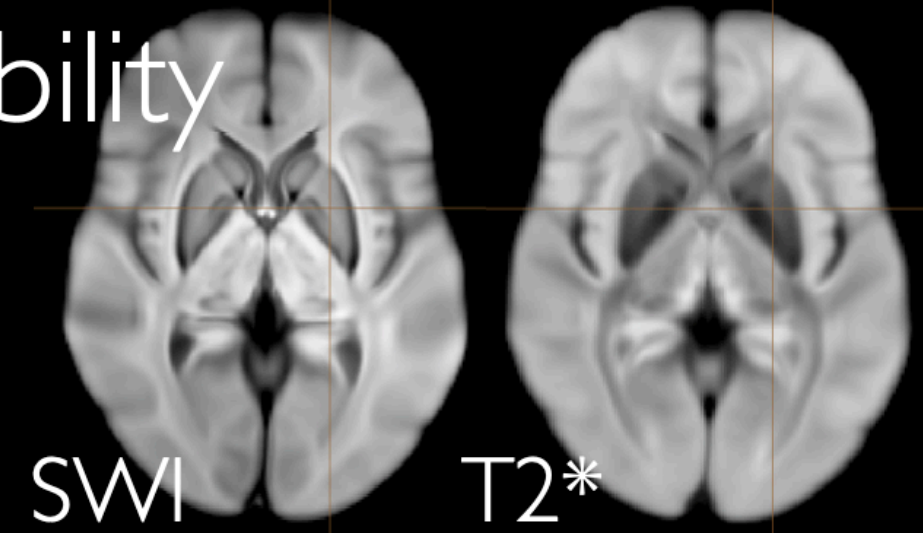
35 mins per subject

biobank^{uk}
Imaging study

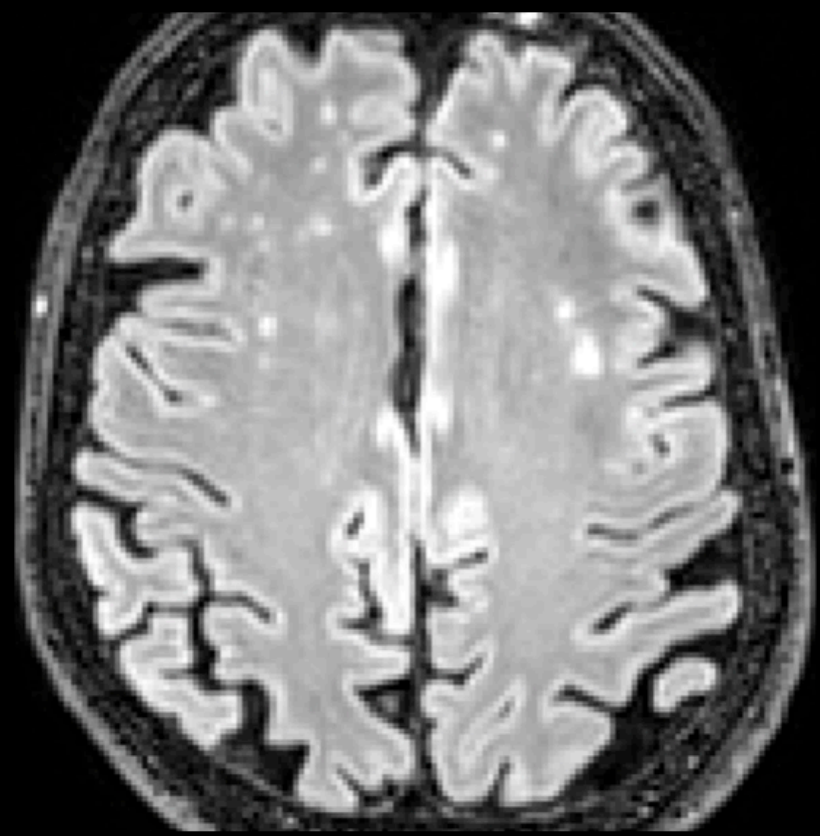
T1-weighted



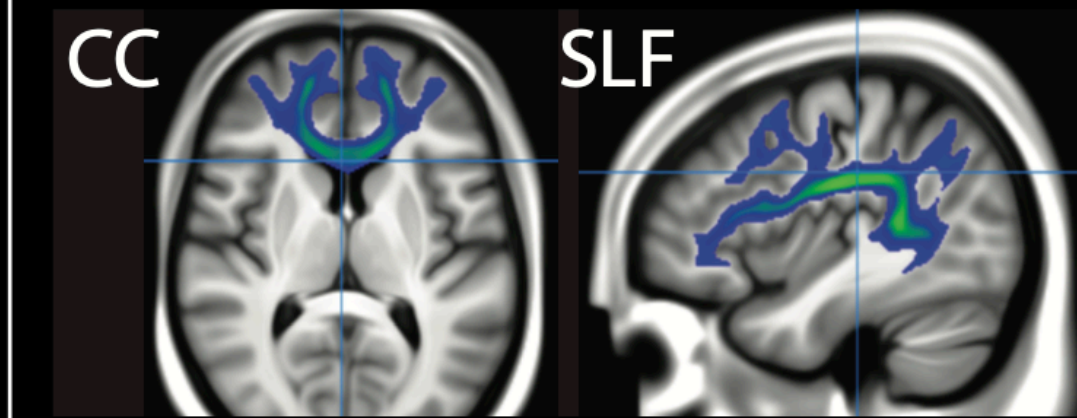
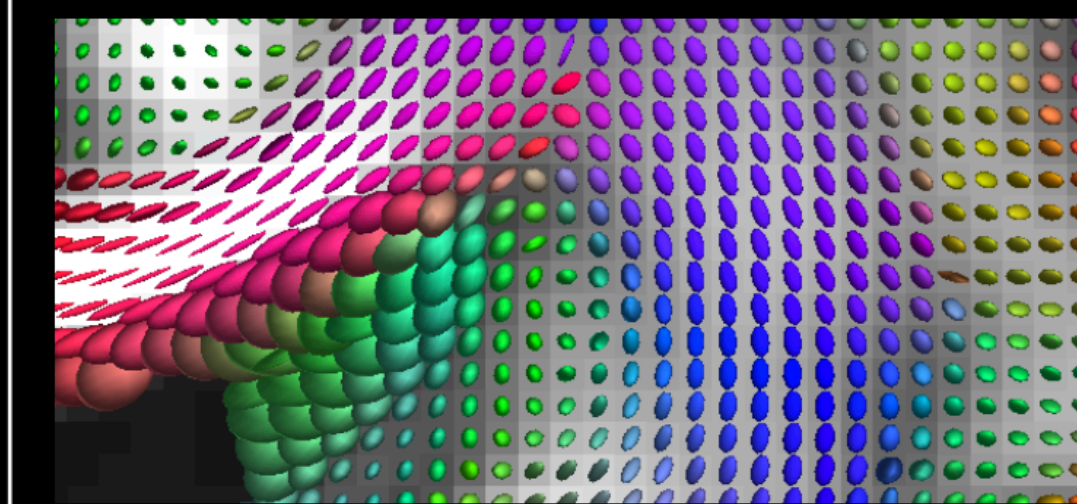
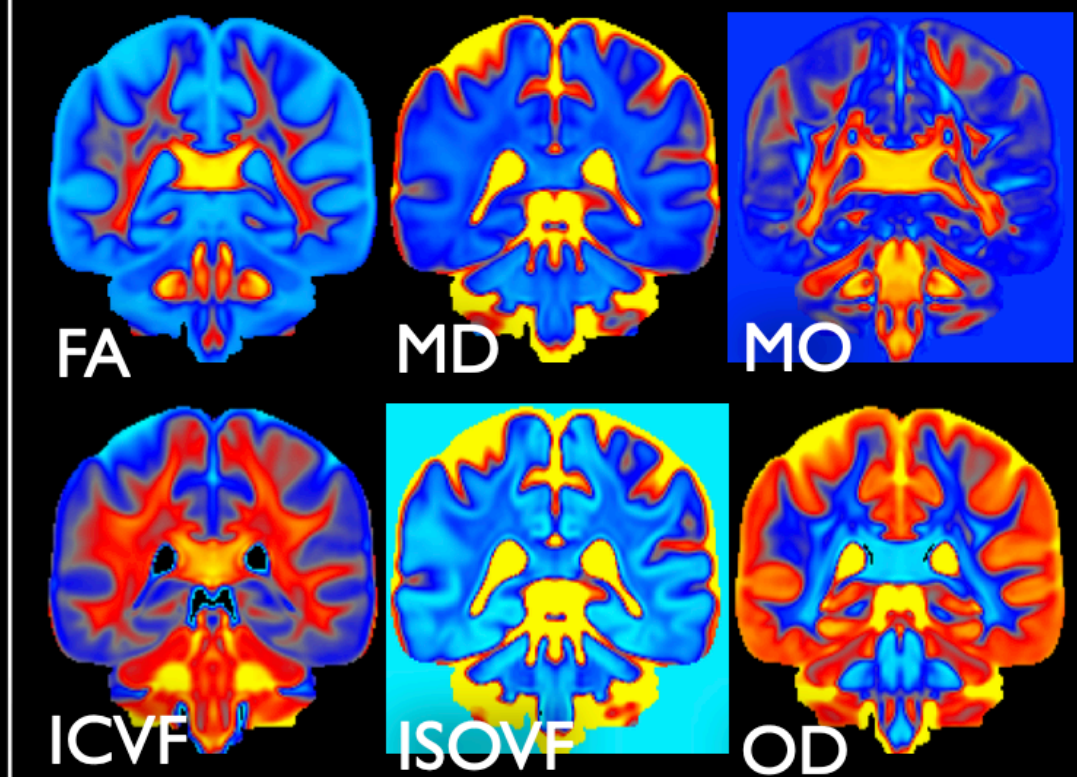
Susceptibility contrast



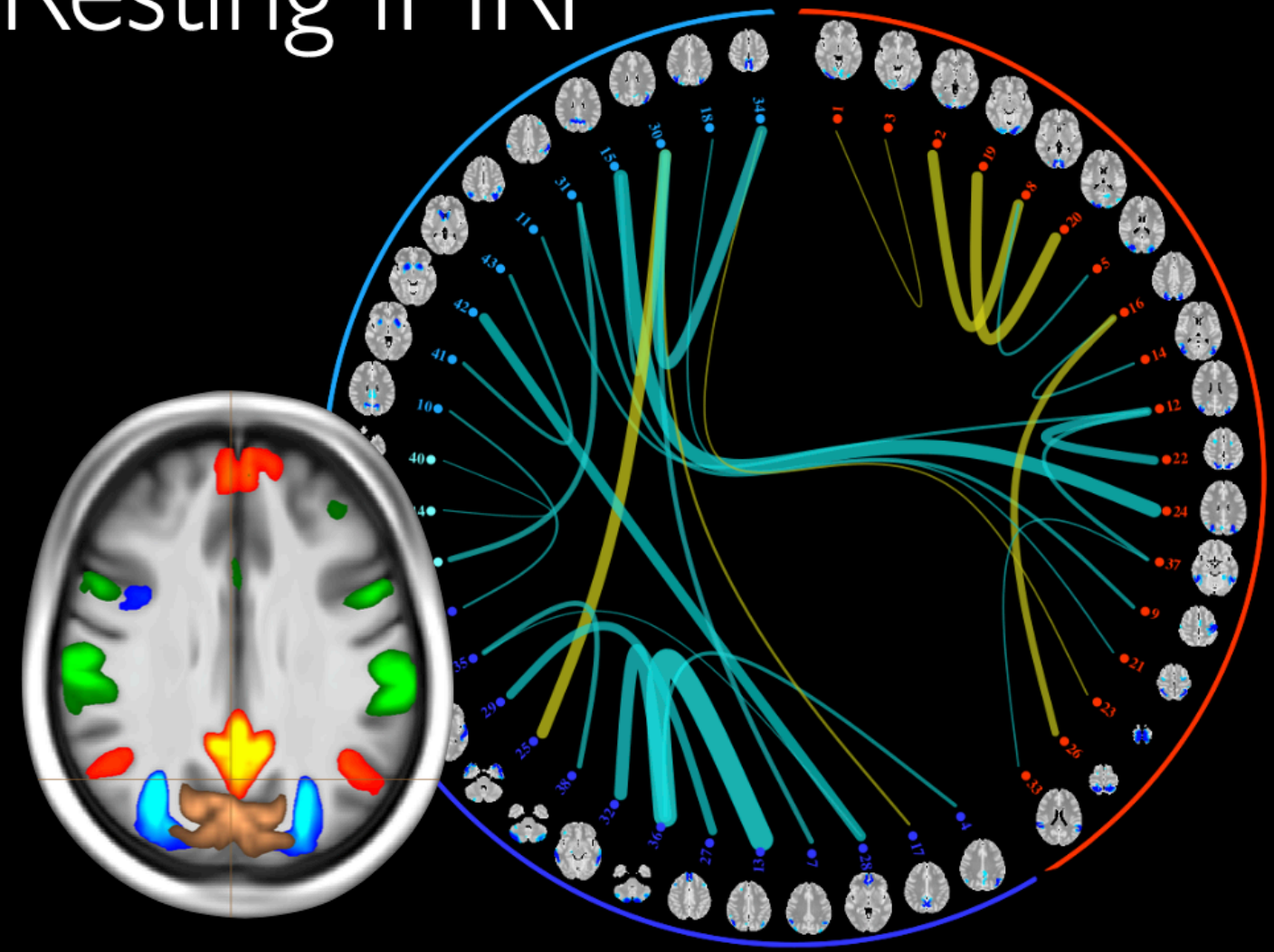
T2 FLAIR



Diffusion MRI



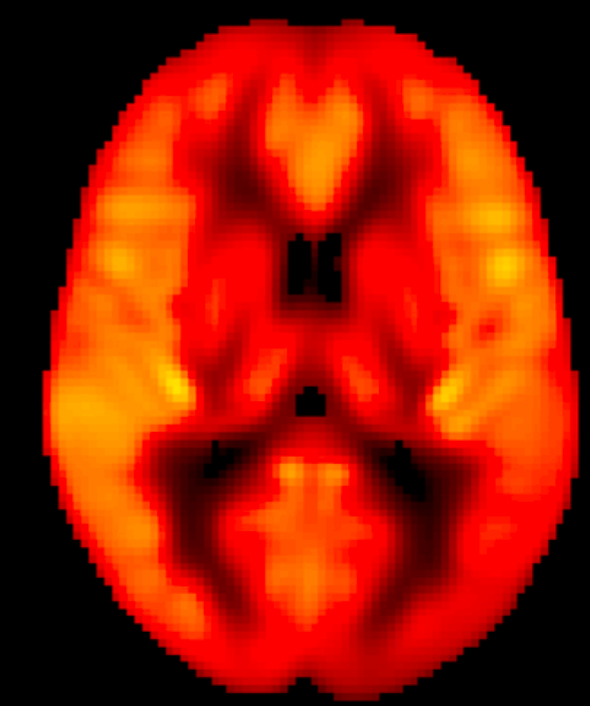
Resting fMRI



Task fMRI



ASL
(repeat scans only)

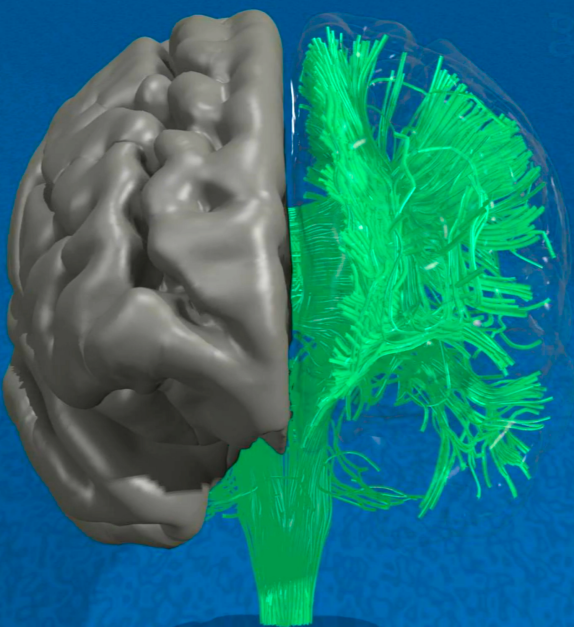


Biobank Neuro Protocol

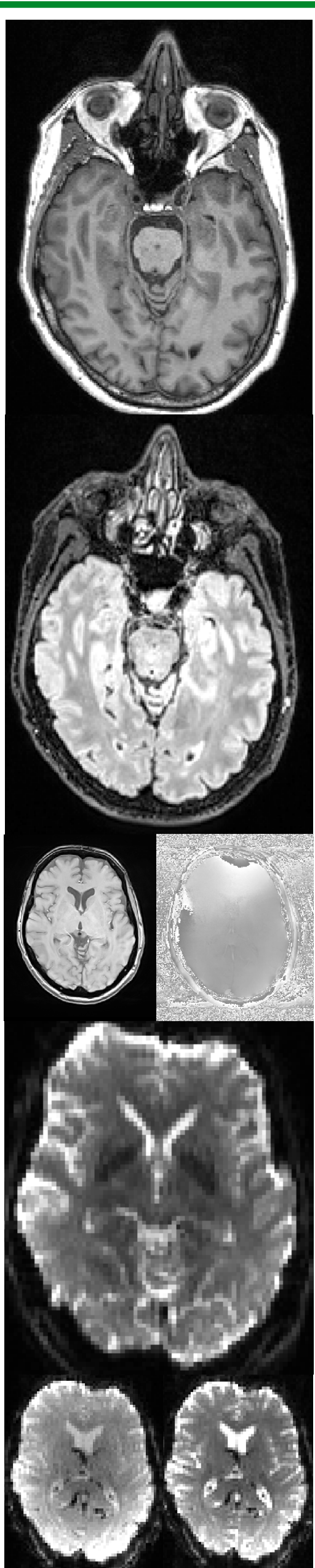
T1	5min	1 mm iso	3D MPRAGE, sagittal, R=2, TI/TR=880/2000ms
rfMRI	6min	2.4 mm iso	EPI MB=x8 TR=0.735 s , TE=39ms, R=1, flip angle=52°, fat sat
tfMRI	4min	as rfMRI	as rfMRI, Hariri faces/shapes “emotion” task
T2 FLAIR	6min	1.05 x 1 x 1 mm	FLAIR, 3D SPACE, sagittal, R=2, PF 7/8, fat sat, TI/TR=1800/5000ms, elliptical
dMRI	7min	2 mm iso	EPI MB=x3, 50x b=1000, 50x b=2000 , TE=92ms, fat sat, topup-fieldmap
swMRI	2.5min	0.8 x 0.8 x 3 mm	3D GRE, axial, R=2, PF 7/8, TE1/TE2/TR=9.4/20/27ms

- 3T scanner, 32-channel head-coil
- fMRI & dMRI: accelerated by UMinn CMRR multiband
- phase-encoding-reversed fieldmap from dMRI (FSL-*topup*)
- dMRI: monopolar encoding (better SNR; eddies fixed by FSL-*eddy*)

(Miller et al, Nature Neuroscience 2016)

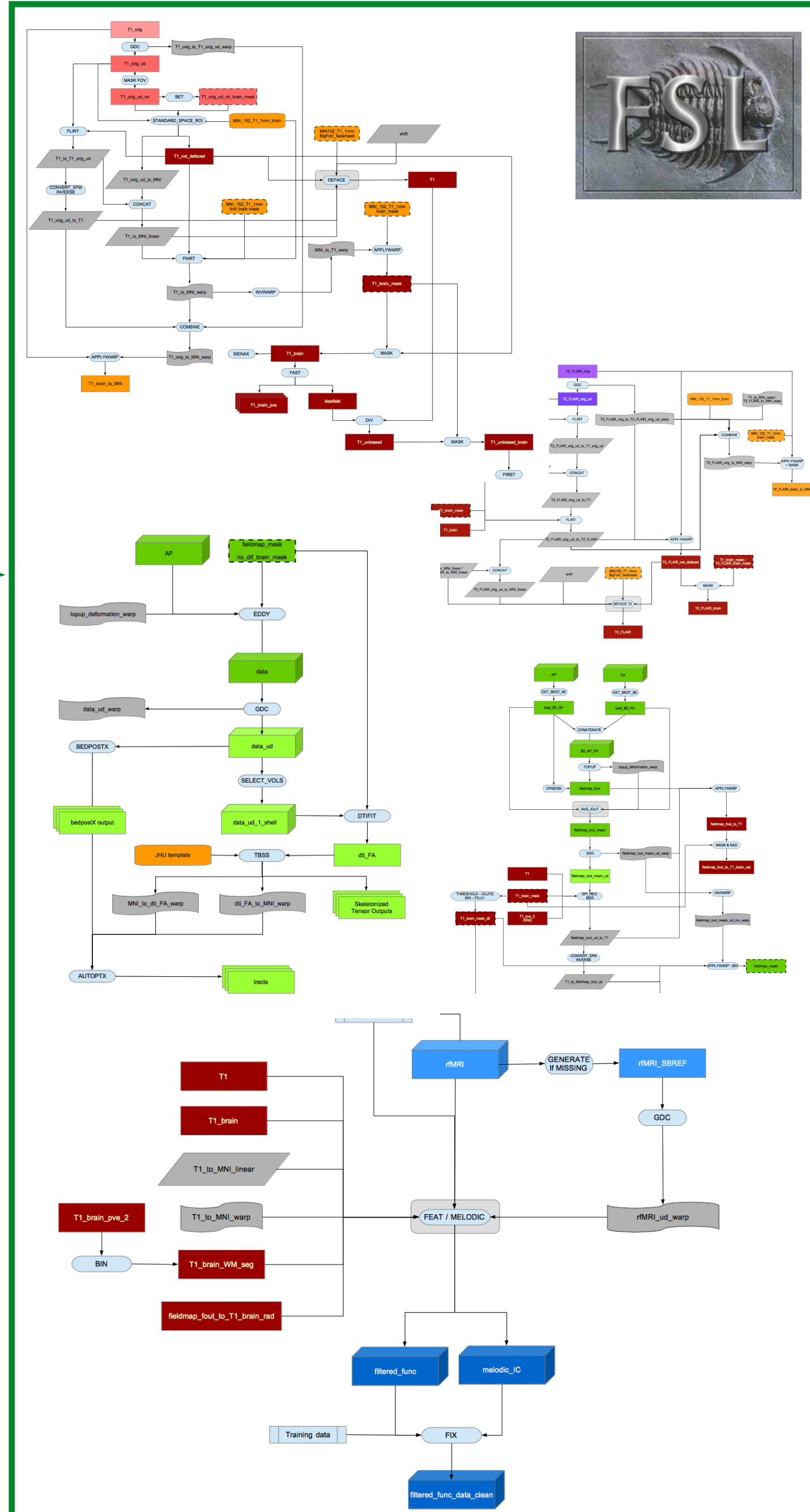


raw data

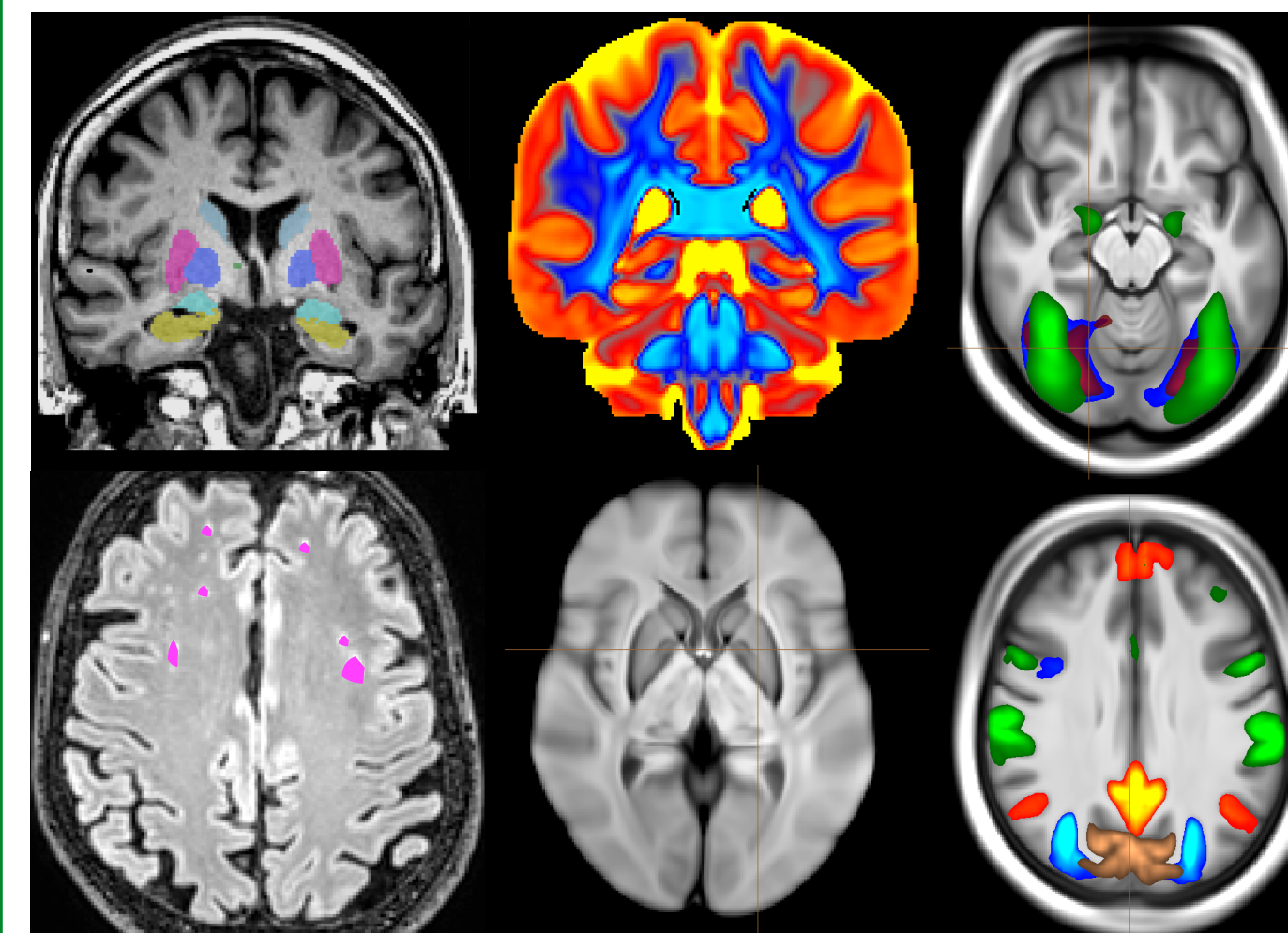


raw DICOMs

automated processing



open-access database



raw and processed NIFTI data

+

imaging-derived phenotypes (IDPs - summary measures)

biobank.ctsu.ox.ac.uk

biobank^{uk}

Index Browse Quick Search Full Search Catalogues Downloads Help

Data-Field 25011
 Description: Volume of thalamus (left)
 Category: T1 structural brain MRI - Brain MRI - Imaging - UK Biobank Assessment Centre

Participants	4,446	Value Type	Continuous, mm3	Sexed	Both sexes
Item count	4,446	Item Type	Data	Instances	Defined (1)
Stability	Accruing	Strata	Derived	Array	No

Data Notes 4 Categories 0 Related Data-Fields 0 Tabulations 0 Resources

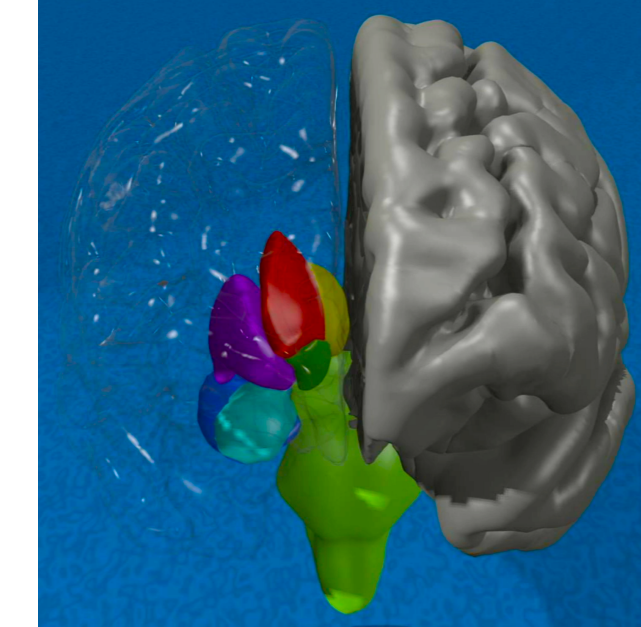
4,446 items of data are available, covering 4,446 participants.
 Defined-instances run from 2 to 2, labelled using Instancing 2.
 Units of measurement are mm3.

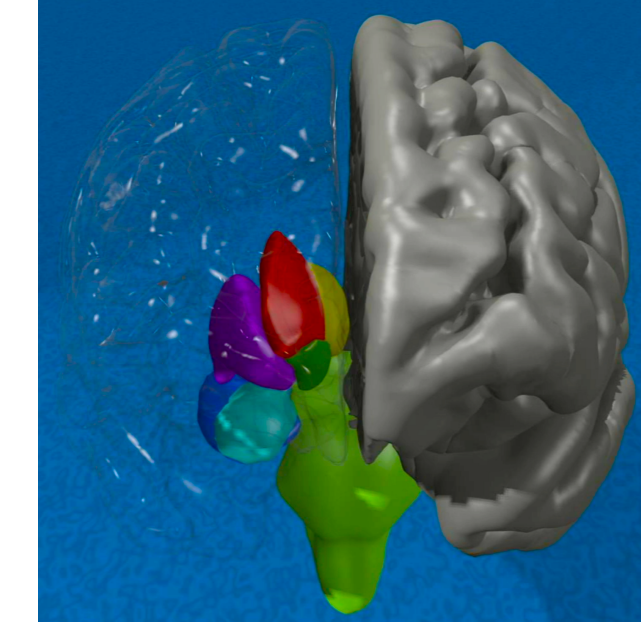
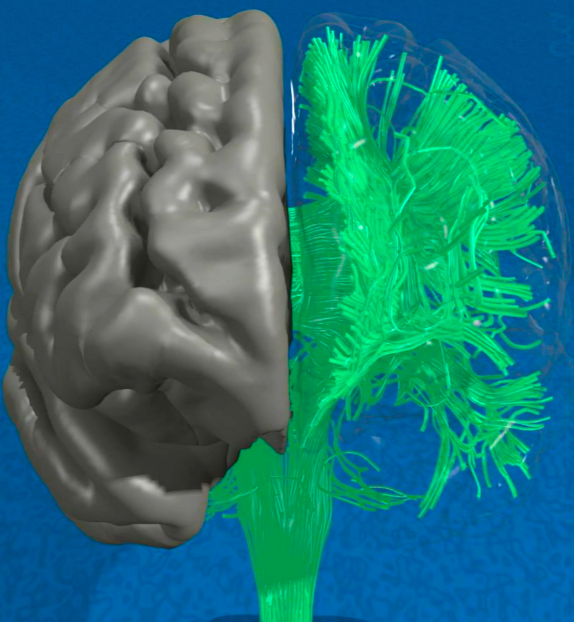
Maximum	11051
Decile 9	8794
Decile 8	8448.5
Decile 7	8205
Decile 6	7989
Median	7781
Decile 4	7601
Decile 3	7408.5
Decile 2	7187
Decile 1	6879
Minimum	4209

• Mean = 7813.07
 • Std.dev = 757.06
 • 1 item below graph minimum of 4964
 • 1 item above graph maximum of 10709

Counts of participants/items last updated 25 Oct 2015.

Improving the health of future generations





tinyurl.com/ukbbrain

biobank^{uk}

Index Browse Quick Search Full Search Catalogues Downloads Help

Data-Field 25011

Description: Volume of thalamus (left)
 Category: T1 structural brain MRI - Brain MRI - Imaging - UK Biobank Assessment Centre

Participants	4,446	Value Type	Continuous, mm3	Sexed	Both sexes
Item count	4,446	Item Type	Data	Instances	Defined (1)
Stability	Accruing	Strata	Derived	Array	No

Data Notes 4 Categories 0 Related Data-Fields 0 Tabulations 0 Resources

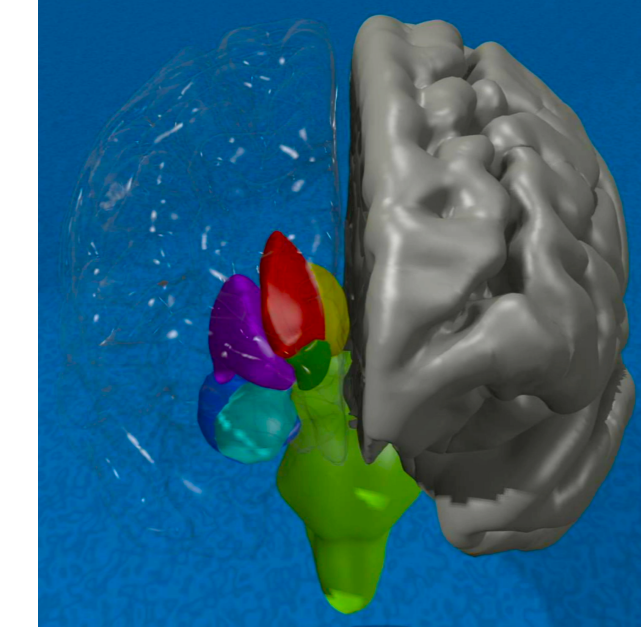
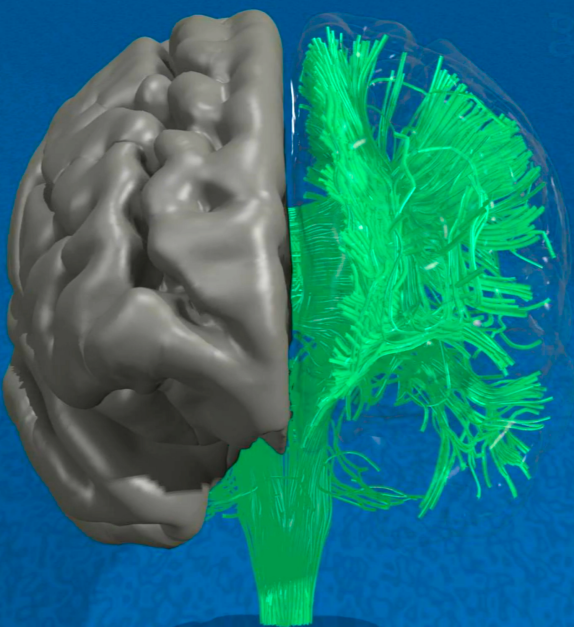
4,446 items of data are available, covering 4,446 participants.
 Defined-instances run from 2 to 2, labelled using Instancing 2.
 Units of measurement are mm3.

Maximum	11051
Decile 9	8794
Decile 8	8448.5
Decile 7	8205
Decile 6	7989
Median	7781
Decile 4	7601
Decile 3	7408.5
Decile 2	7187
Decile 1	6879
Minimum	4209

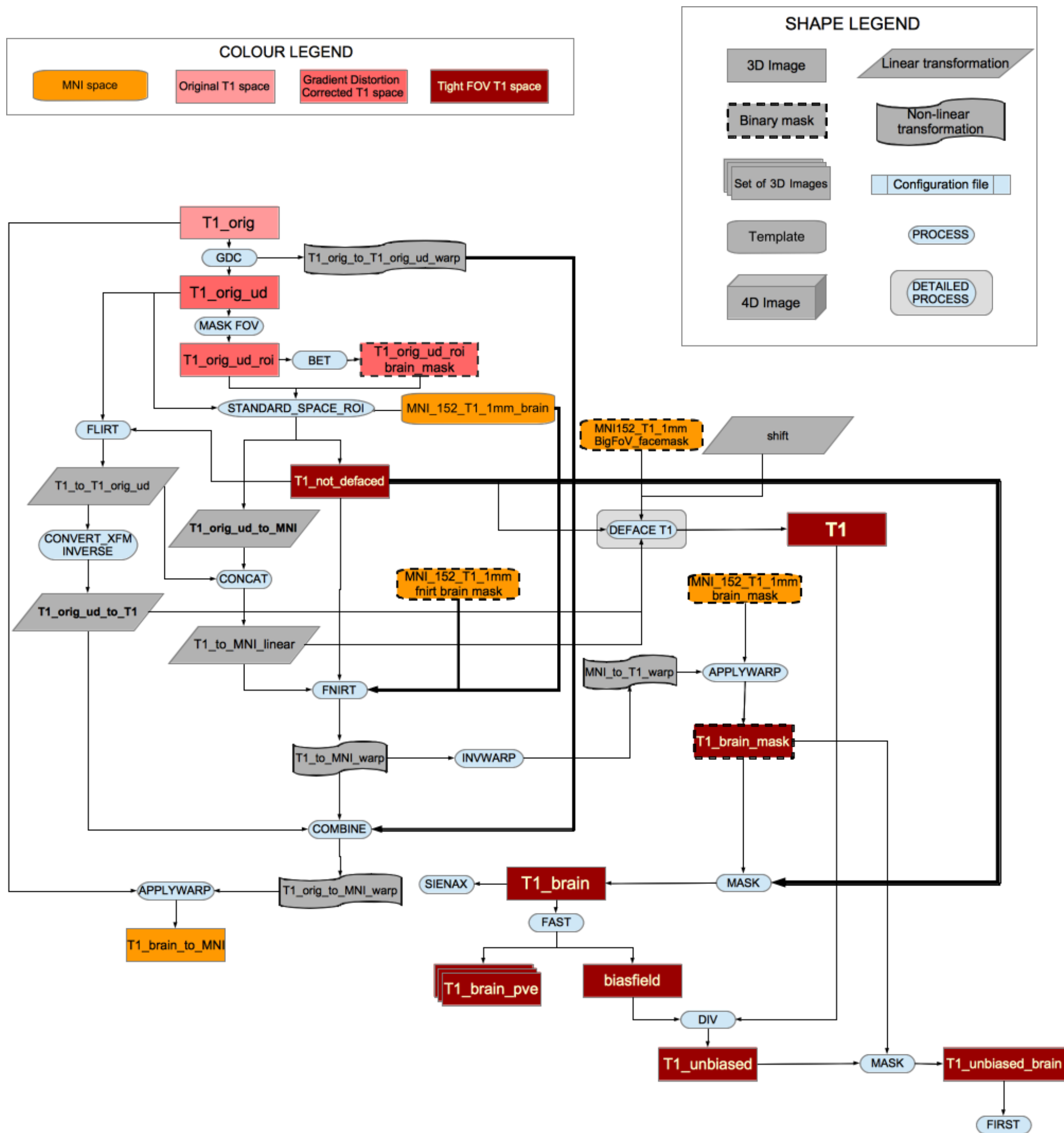
- Mean = 7813.07
- Std.dev = 757.06
- 1 item below graph minimum of 4964
- 1 item above graph maximum of 10709

Counts of participants/items last updated 25 Oct 2015.

Improving the health of future generations

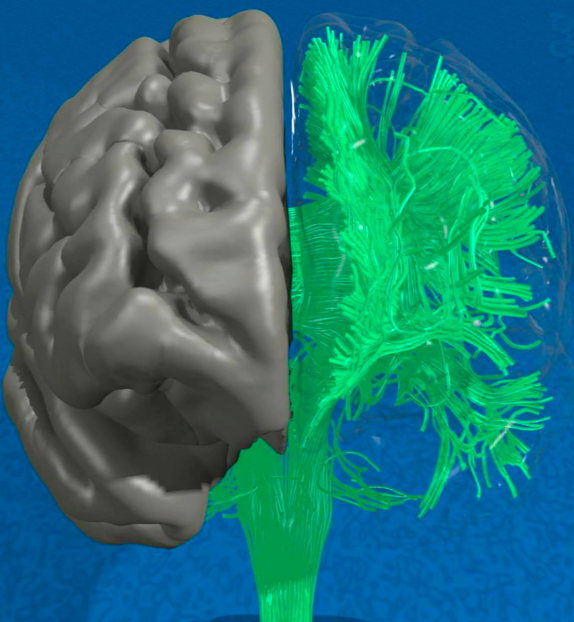


(Alfaro-Almagro et al, NeuroImage, 2018)



- **Processing pipeline:** Fully automated computing structure.
- Basic blocks are neuroimaging tools (e.g. FSL): Process the images and generate a new output image or value that can be used in the next step.
- The pipeline structure is mainly written in **bash** and **python**.
- The basic blocks that are not FSL are tools written in Matlab, R, Java, etc. Some of them belong to software libraries such as FreeSurfer.

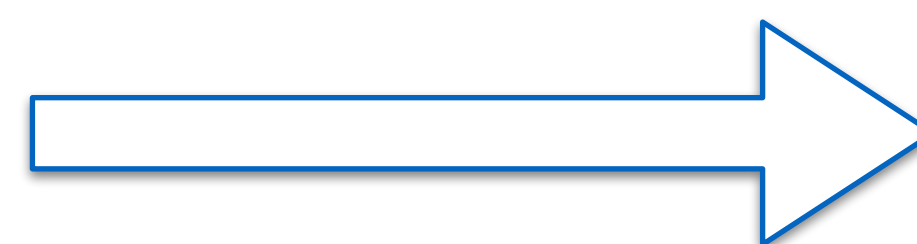
* FMRIB Software Library



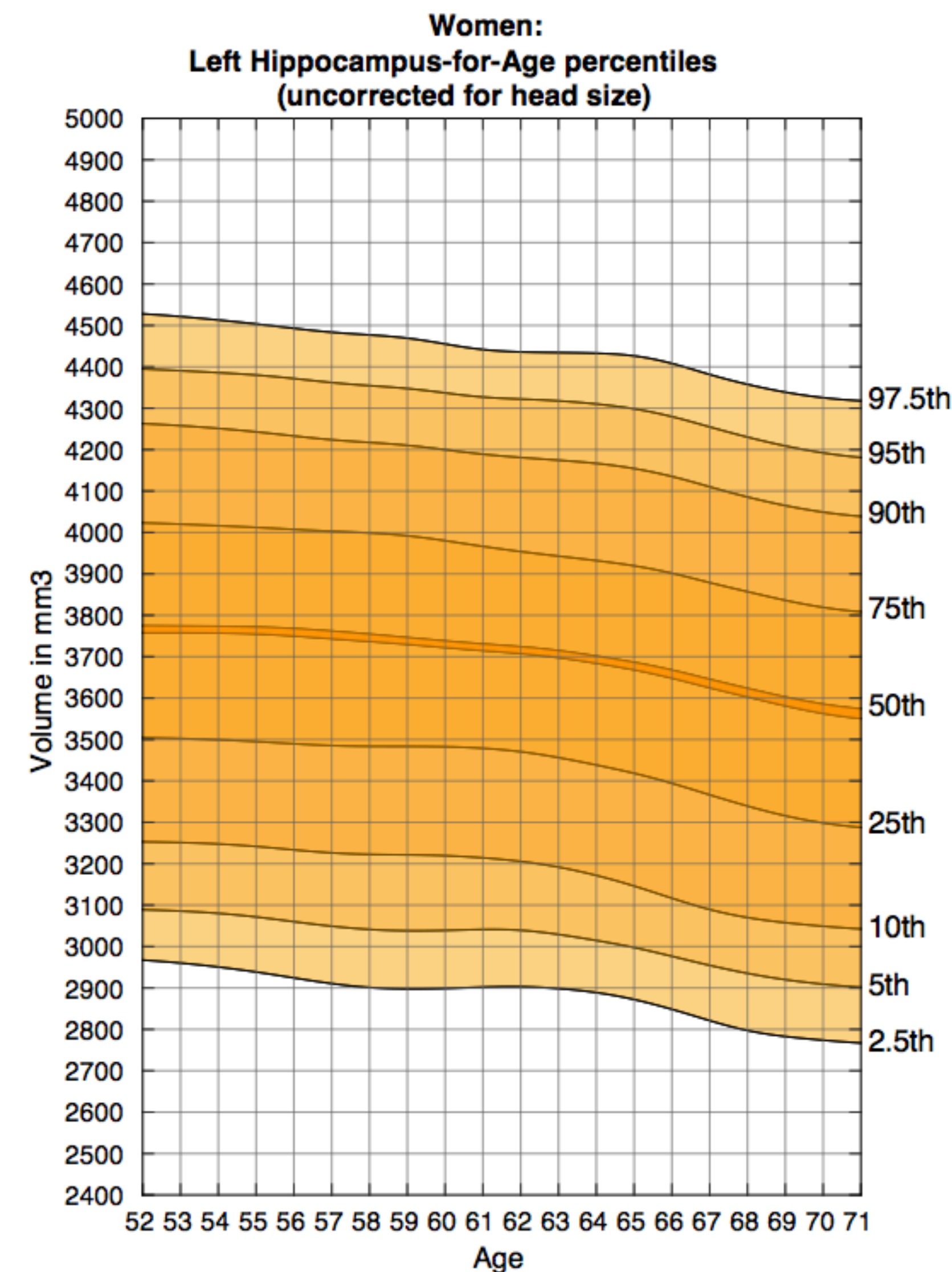
FSL tools in the pipeline

Nomograms with normative percentiles of subcortical volumes derived from over 19,700 people in UK Biobank

FSL Tool	Number of IDPs
FIRST	29
SIENAX	11
FLIRT + FNIRT + FAST + VBM	139
Bianca	3
FEAT	16
dtifit + eddy + TBSS	432
dtifit + eddy + Tractography	243
FIX + Melodic + FSLNets	3472

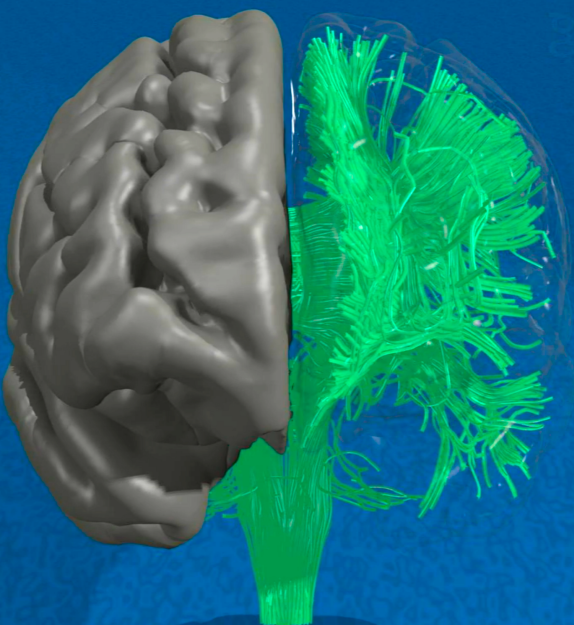


Cool clinical papers!

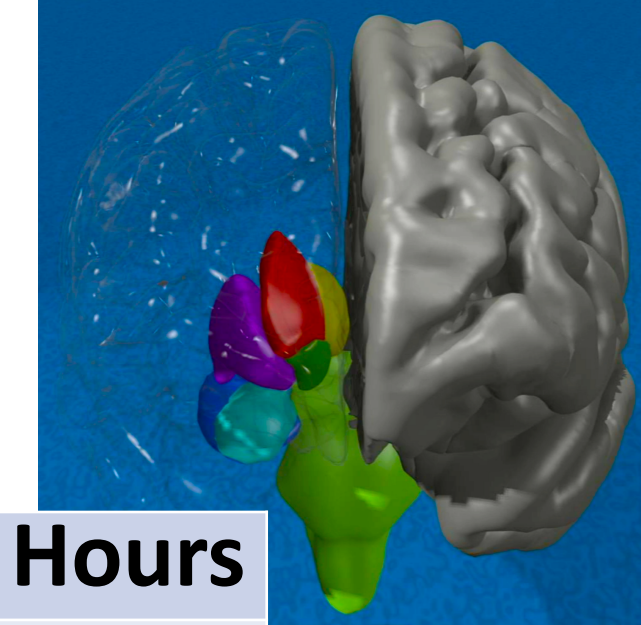


(Nobis et al, NeuroImage: Clinical, 2019)





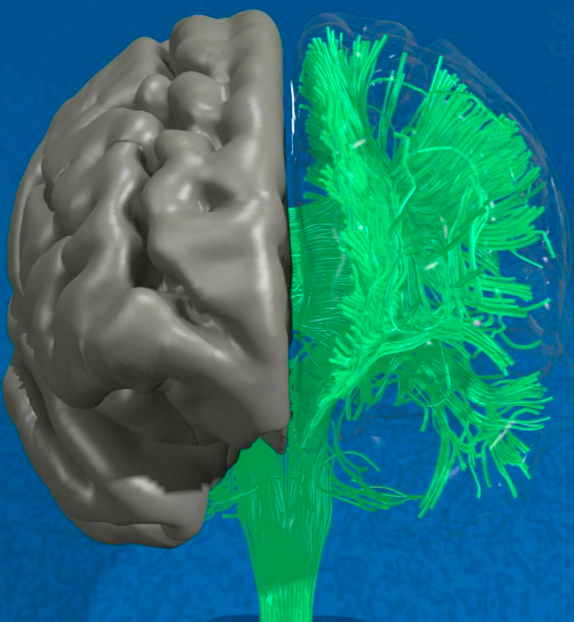
Big Data Needs for the pipeline



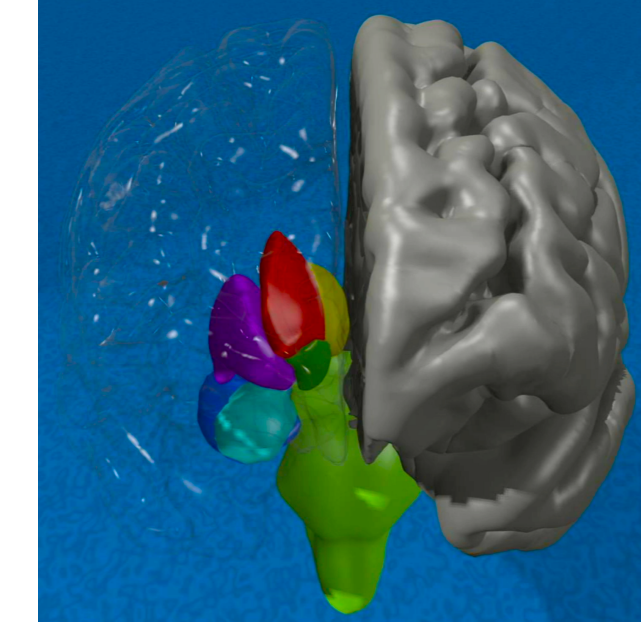
- ~6 GB per subject = ~1 PB total data
- ~21 CPU hours and 1 GPU hours per subject. Some of the tasks may run in parallel, reducing it to ~12 CPU hours.
- With no GPUs, ~60 CPU hours per subject
- Imaging researchers may run their own analyses in the cloud
- Easy to run on DNA Nexus - RAP
- Repo: https://git.fmrib.ox.ac.uk/falmagro/uk_biobank_pipeline_v_1.5
- Docker: <https://git.fmrib.ox.ac.uk/paulmc/fbp-dockerfile>
- More resources: <https://www.fmrib.ox.ac.uk/ukbiobank/>
- For any question: fidel.alfaroalmagro@ndcn.ox.ac.uk

CPU Job	Hours
T1 + T2 FLAIR proc	2.5
swMRI proc	0.2
QSM	0.4
ASL	0.6
Topup	0.3
dMRI proc	0.2
TBSS	0.5
NODDI	0.2
Melodic	2.1
FIX	0.5
Feat	1
FreeSurfer	12
IDP Generation	0.15
Total CPU Hours	20.65

GPU Job	Mins
EDDY	18
bedpostx	9
autoPtx	29
Total GPU Hours	~1



biobank^{uk}



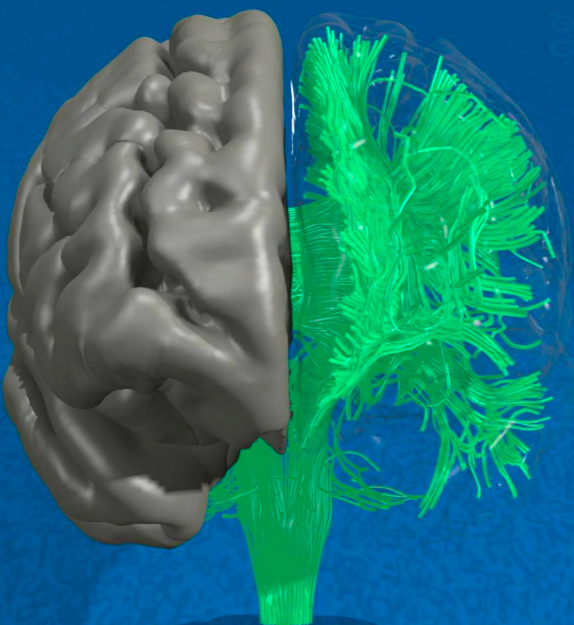
1. Background
2. Brain Imaging Project
3. **Brain Imaging: What can we do?**



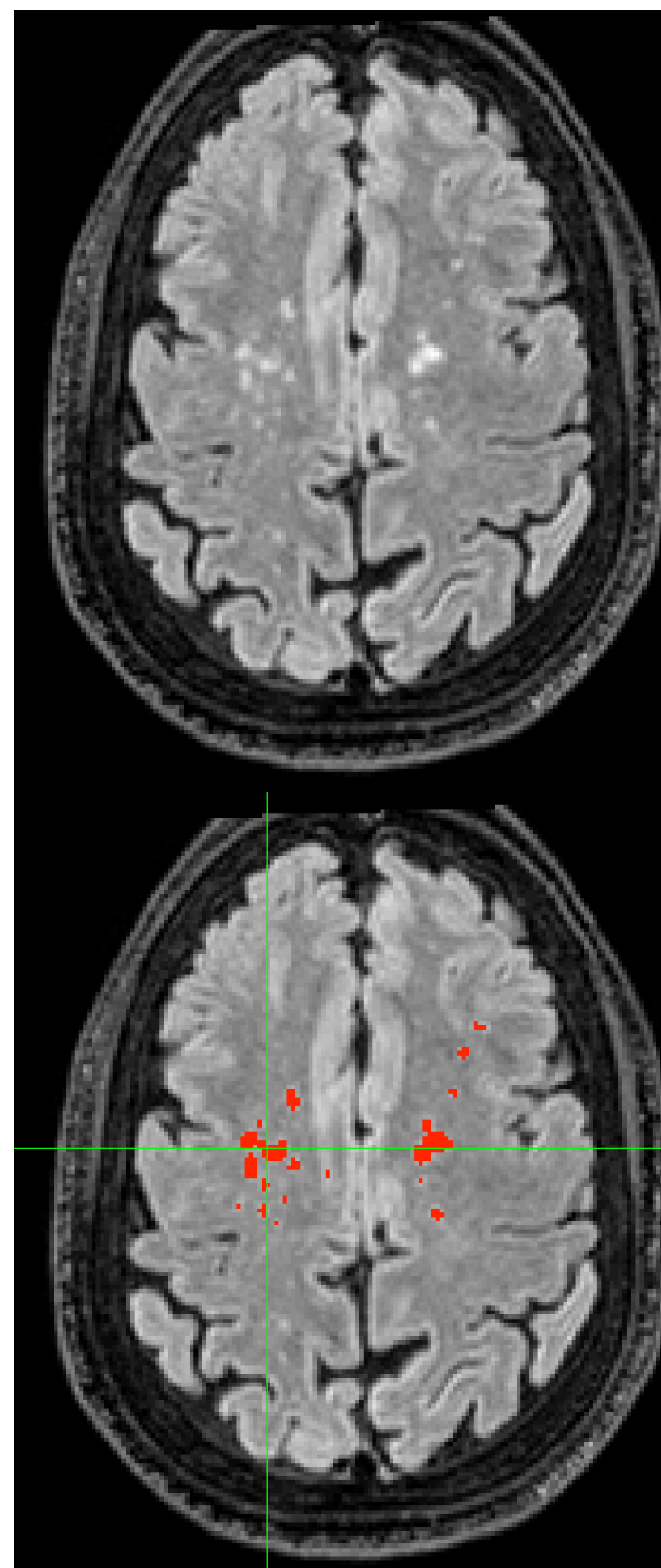
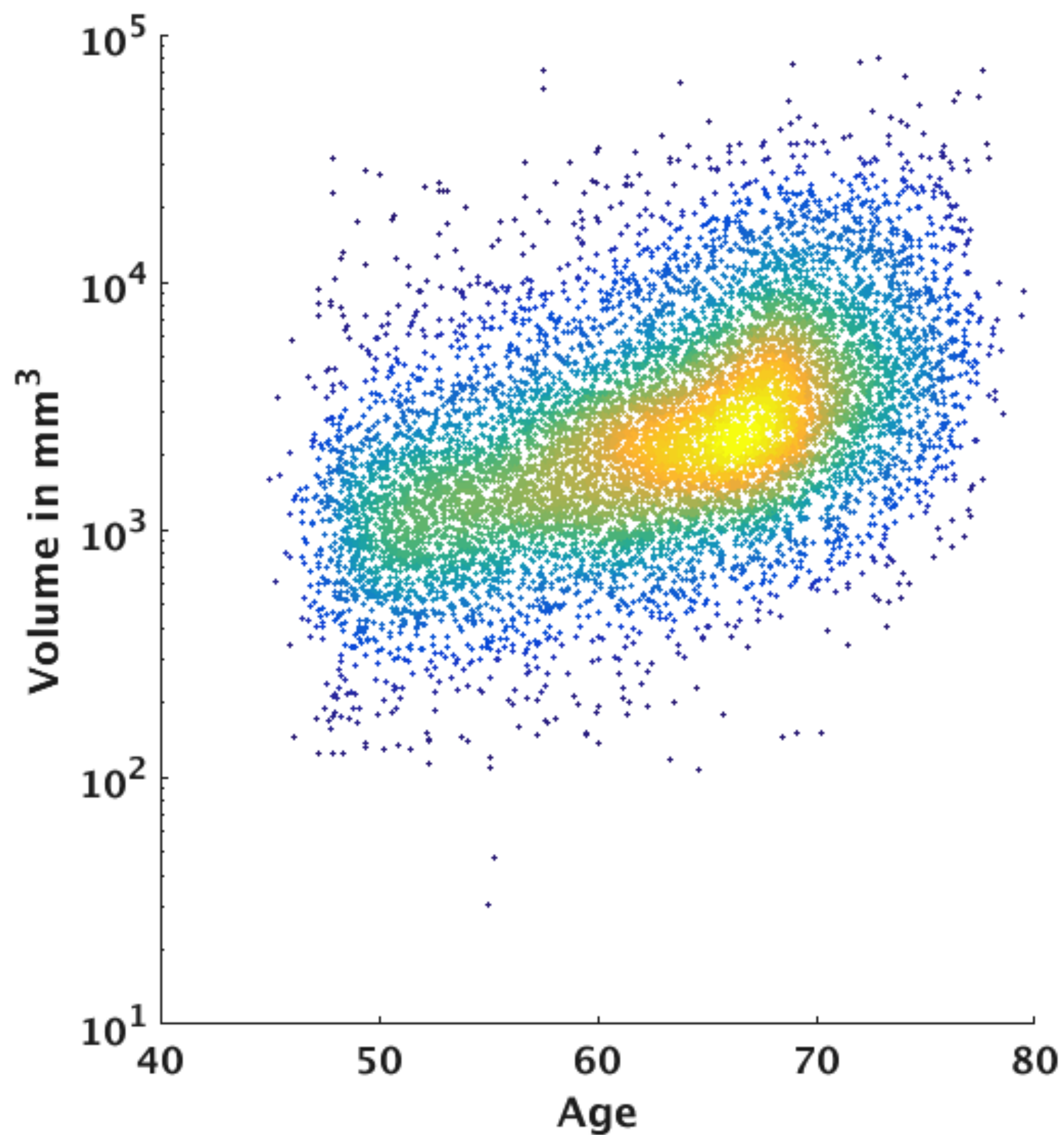
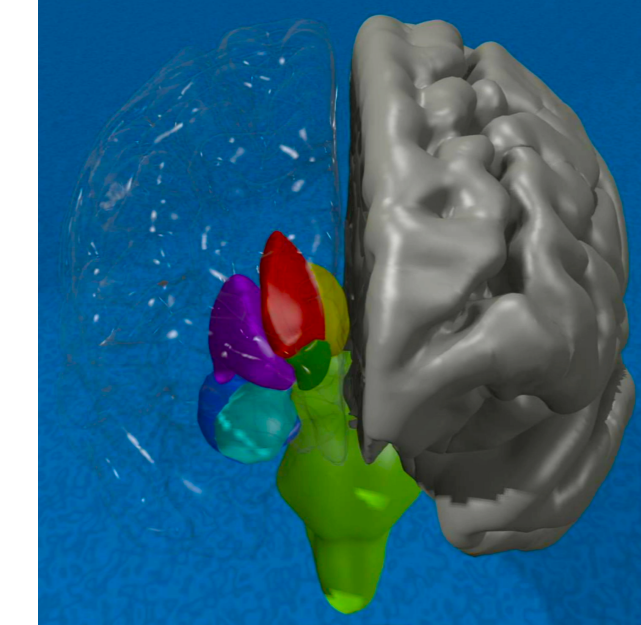
UNIVERSITY OF
OXFORD



wellcome
centre
integrative
neuroimaging

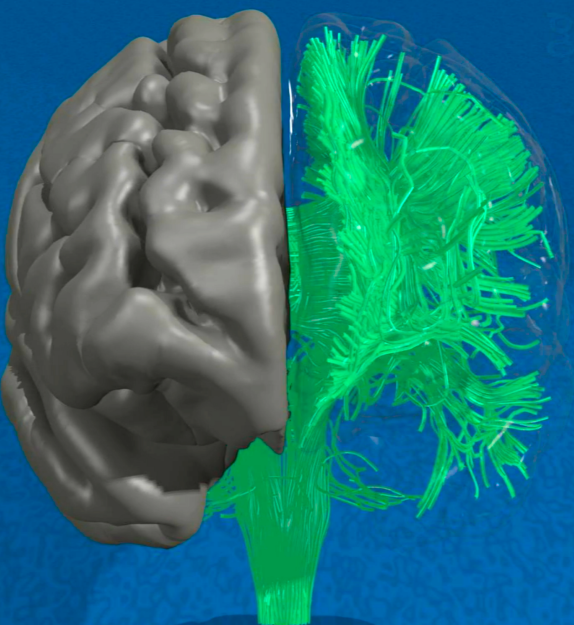


T2 FLAIR lesions: Correlation with age



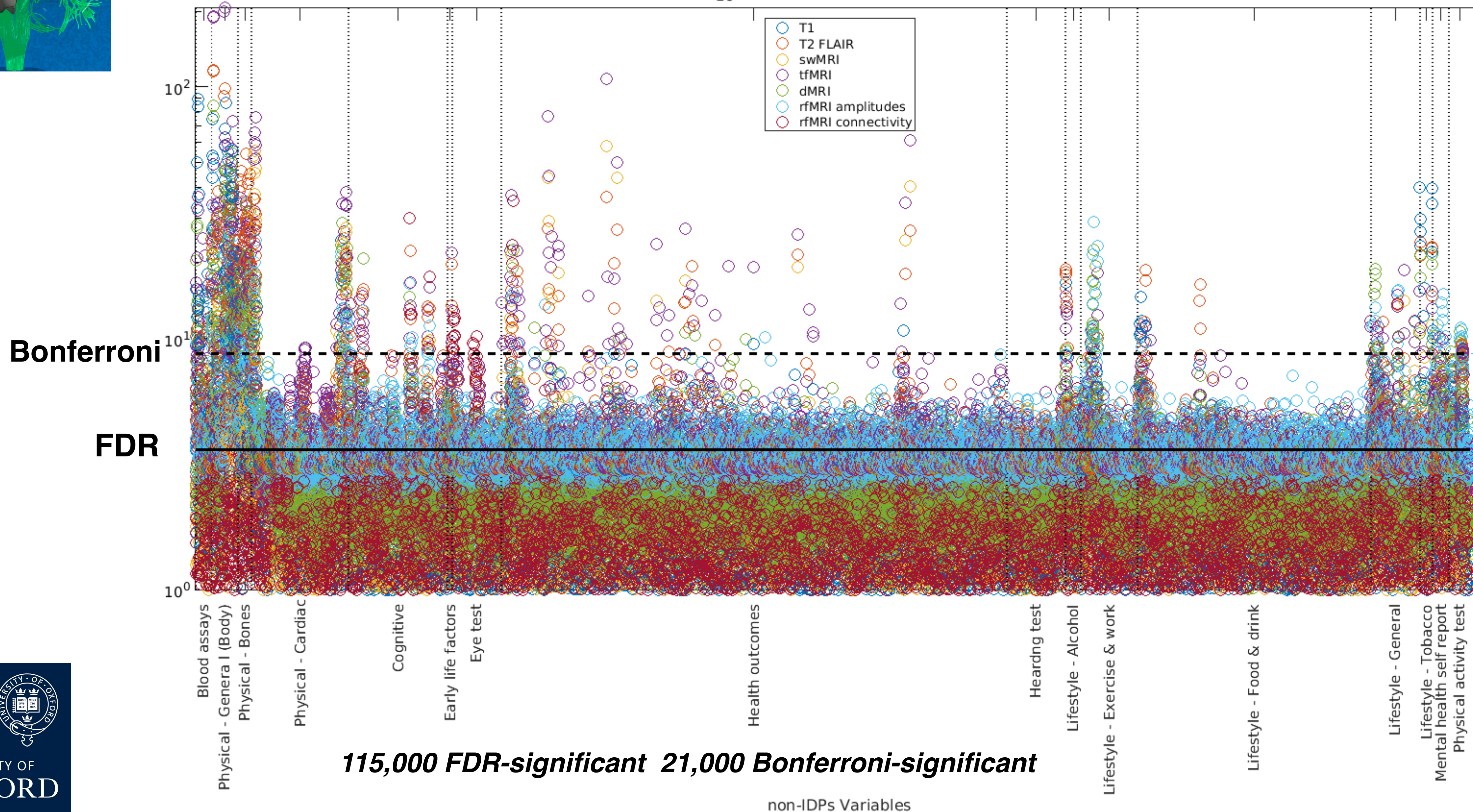
BIANCA

(Griffanti et al, NeuroImage 2018)



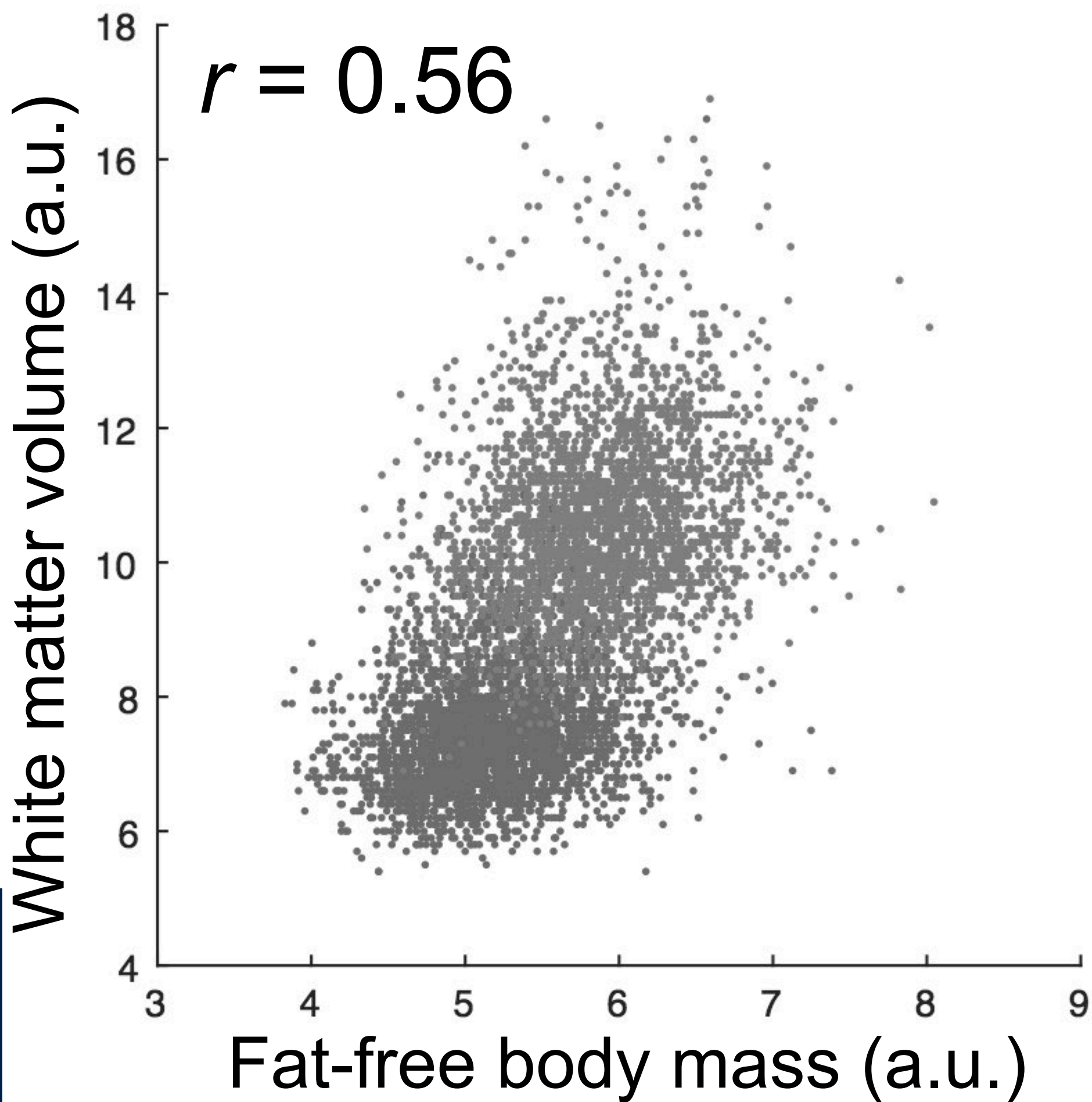
23.5 million univariate associations 2,641 IDPs x 8,907 non-IDPs (21,400 subjects)

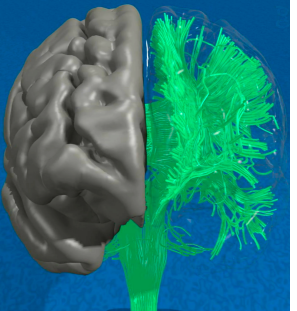
$-\log_{10}$ (P values) full unconfounding



Danger of ignoring confounds

- Simple associations affected dramatically by shared confounds
- Including inappropriate pooling across groups

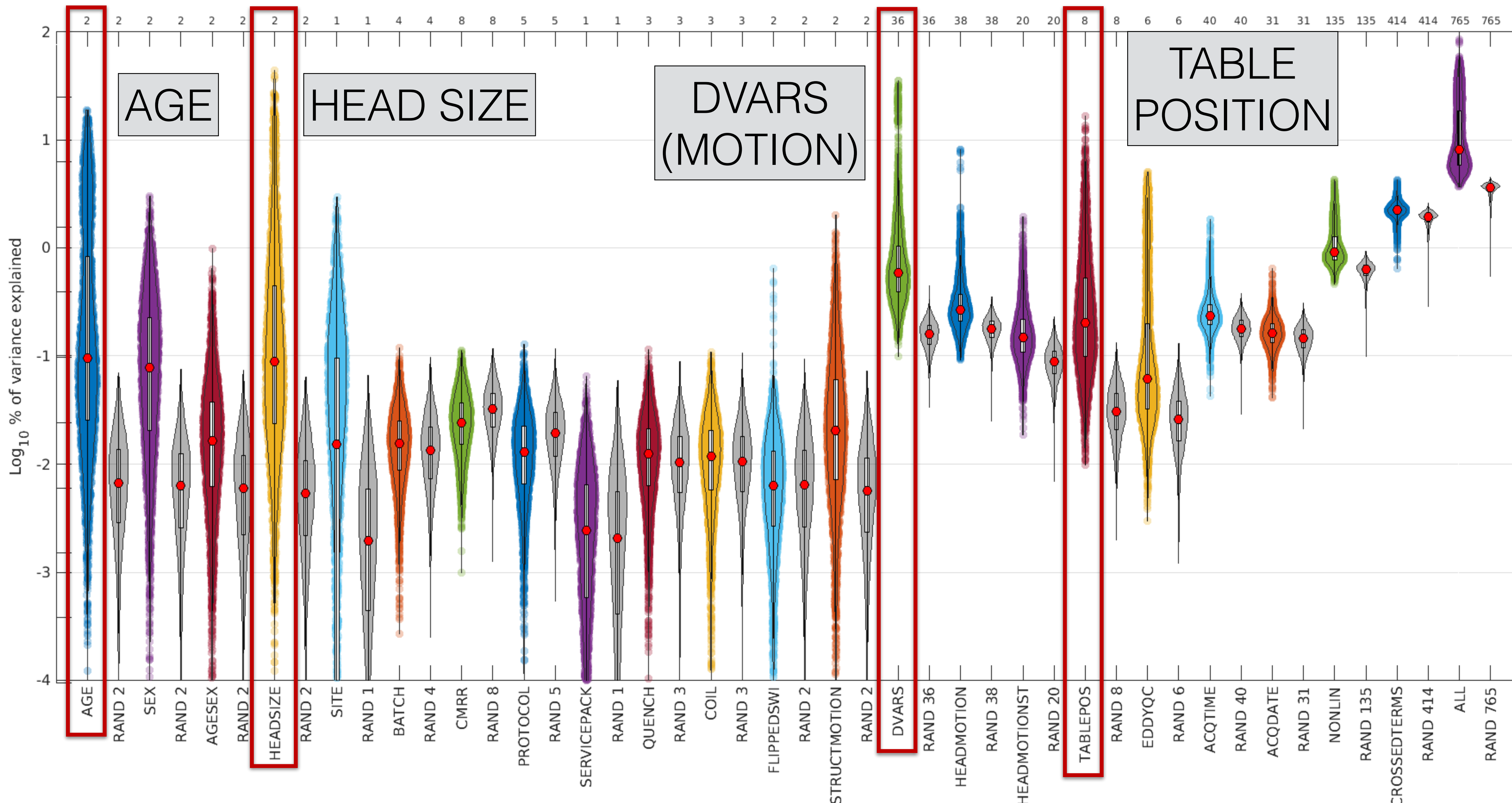


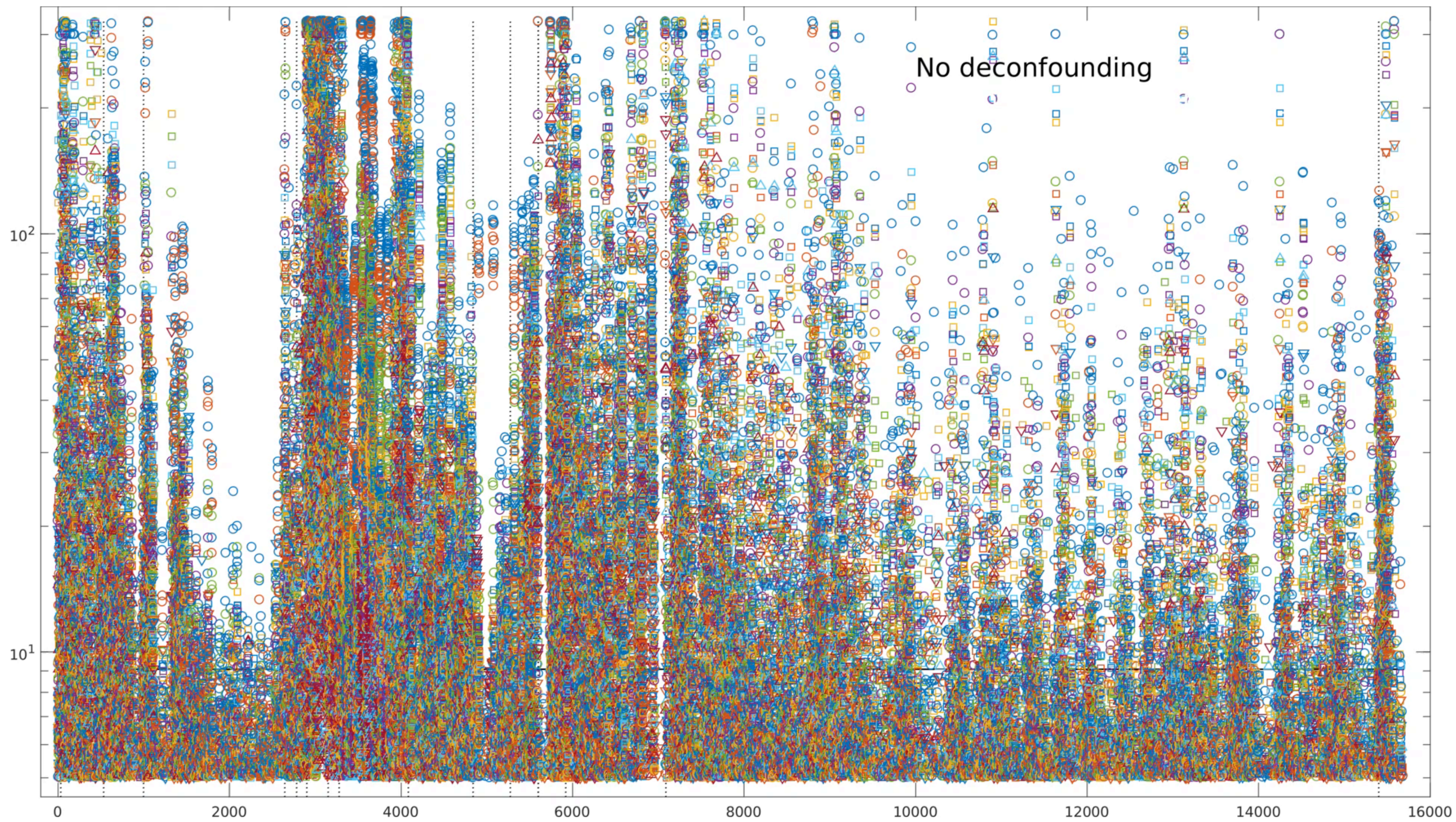


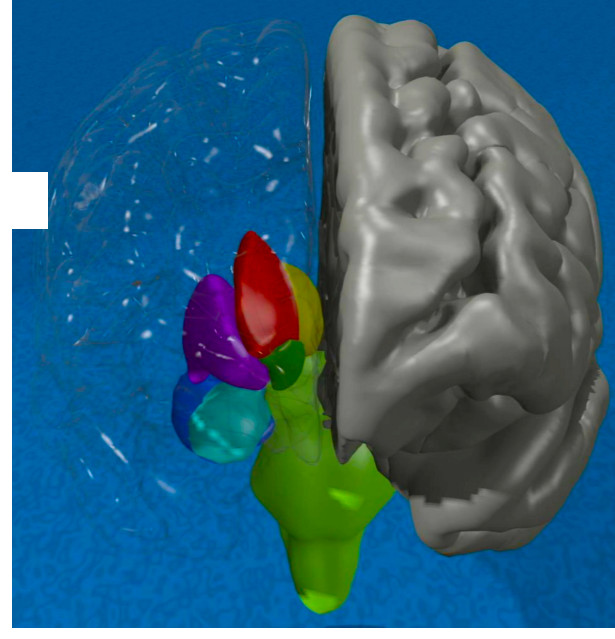
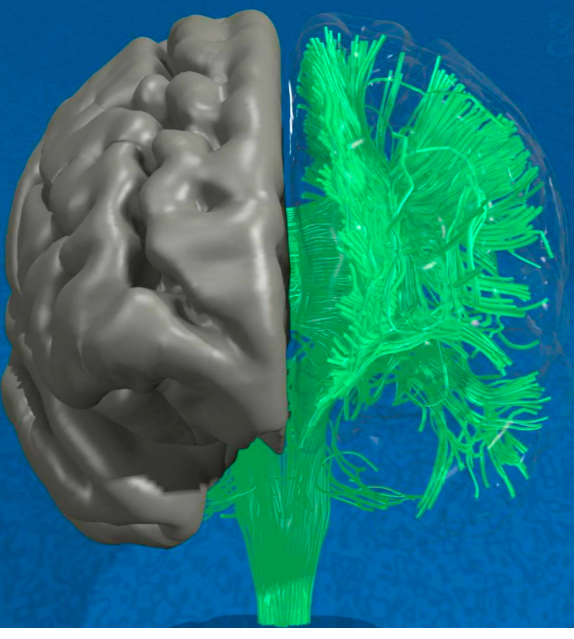
Confounds in imaging

(Alfaro-Almagro et al, NeuroImage, 2020)

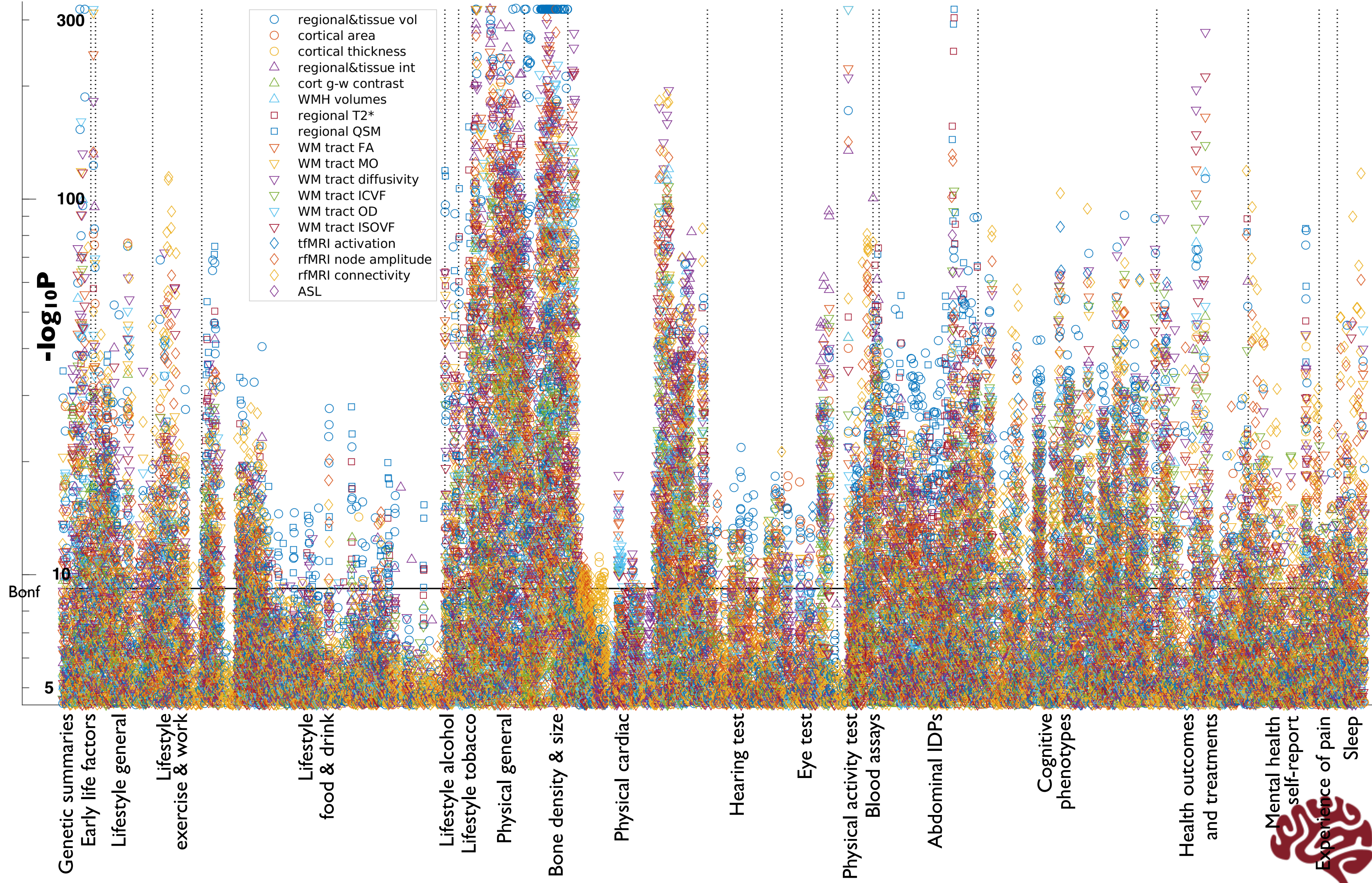
ALL IDPs Unique Variance Explained

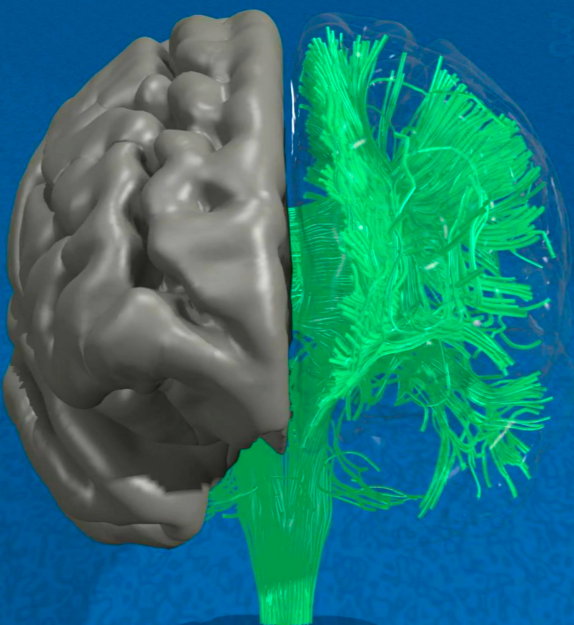




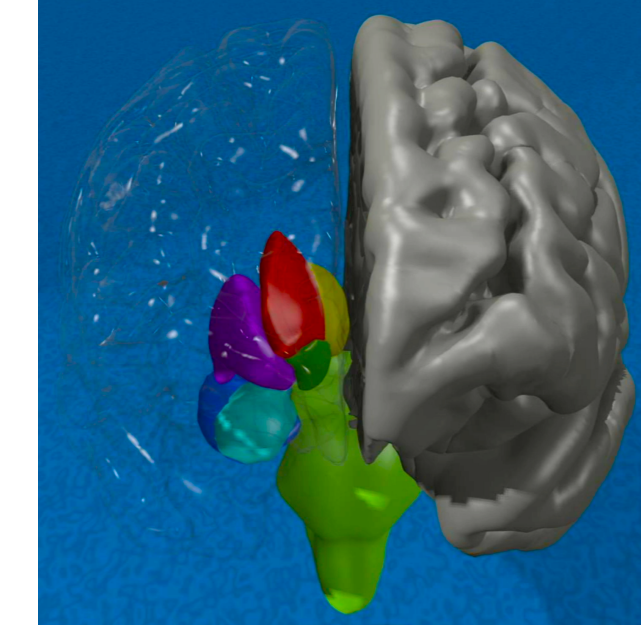


- 82k UKB subjects
- 2.1m FDR-significant, 324k Bonferroni-significant



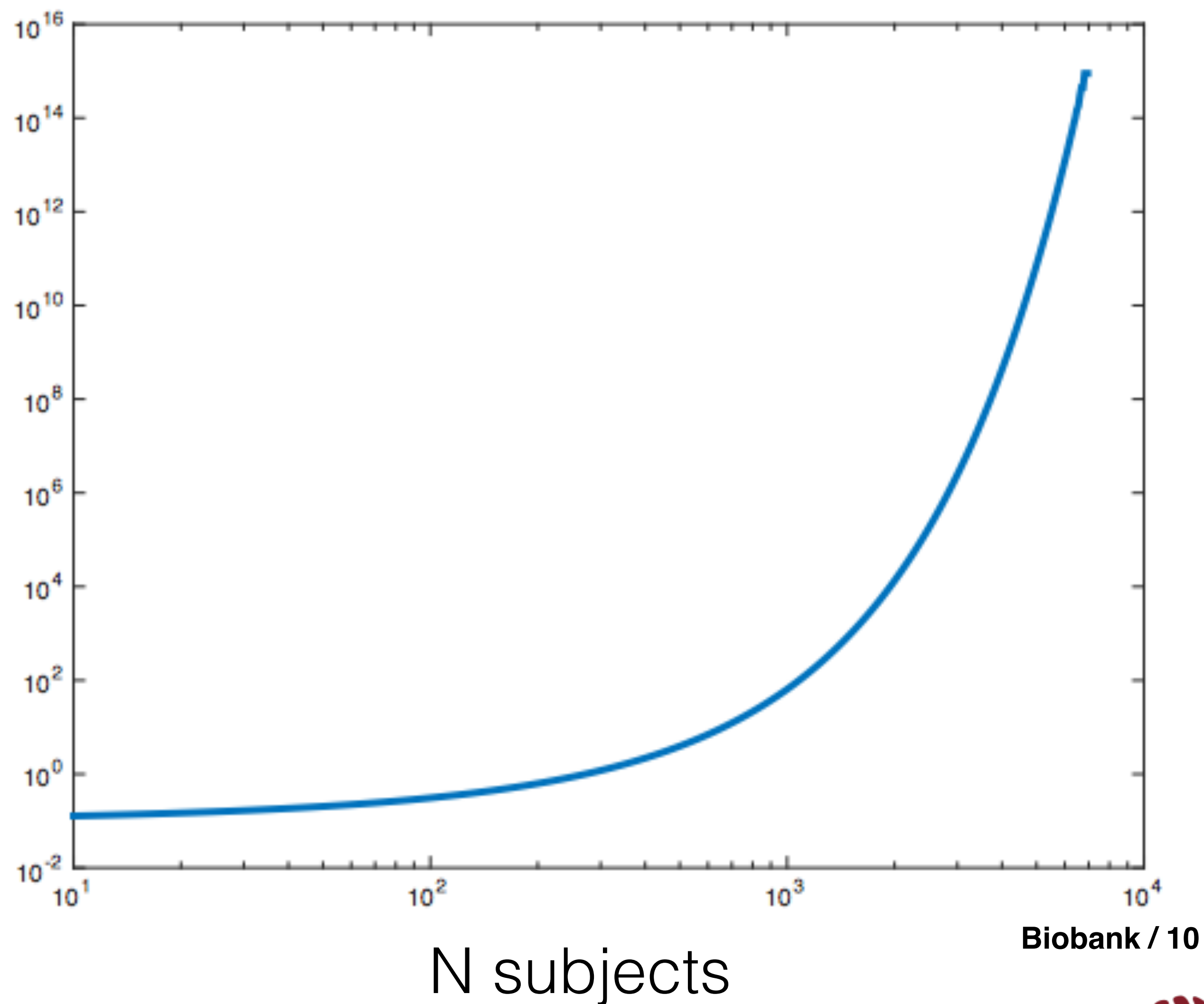


stats of Big Data

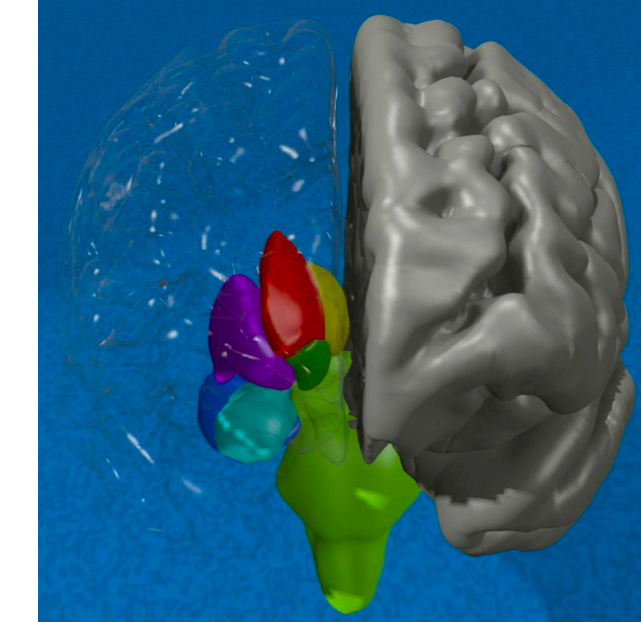
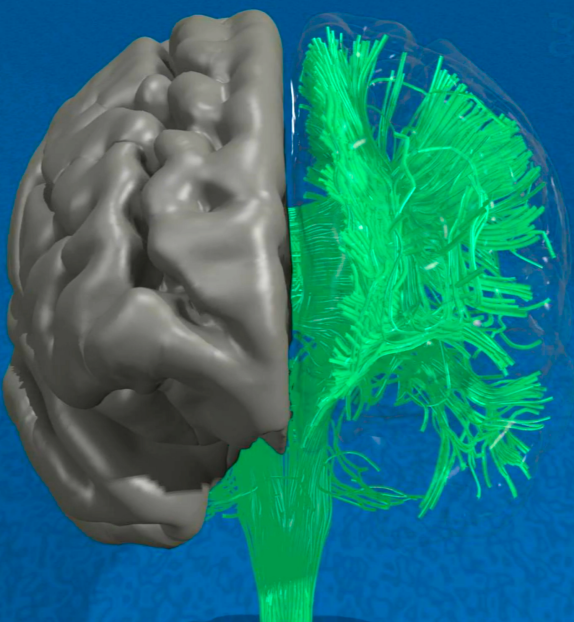


Number of different tests
passing Bonferroni
with very low correlation
given $r=0.1$

Probably better if we start
reporting more size effects
and power in correlations?



(Steve Smith & Tom Nichols, *Neuron*, 2018)



MERCI BEAUCOUP!

Brain Imaging Contributors

Scientific direction Stephen Smith, Karla Miller (FMRIB, Oxford)

Processing pipeline Fidel Alfaro-Almagro (FMRIB, Oxford)

Additional input on acquisitions/protocols/reconstruction/processing Neal Bangerter (Brigham Young), Kamil Ugurbil, Essa Yacoub, Steen Moeller, Eddie Auerbach (CMRR, U Minnesota), Junqian Gordon Xu (Mount Sinai), David Thomas, Daniel Alexander, Gary Zhang, Enrico Kaden (UCL), Alessandro Daducci (EPFL), Tony Stoecker (Rhineland Study/Bonn), Stuart Clare, Heidi Johansen-Berg (FMRIB, Oxford), Deanna Barch, Greg Burgess, Nick Bloom, Dan Nolan, Michael Harms, Matt Glasser (Washington U), Doug Greve, Bruce Fischl, Jonathan Polimeni (MGH), Andreas Bartsch (Heidelberg), Anna Murphy (Manchester), Fred Barkhof (VU Amsterdam/UCL), Christian Beckmann (Donders Nijmegen), Chris Rorden (U South Carolina), Peter Weale, Iulius Dragonu (Siemens UK), Steve Garratt (Project Manager, UK Biobank Imaging)

Additional input on image processing pipeline Mark Jenkinson, Jesper Andersson, Stamatios Sotiropoulos, **Saad Jbabdi**, Ludovica Griffanti, Gwenaelle Douaud, Eugene Duff, Moises Hernandez Fernandez, Emmanuel Vallee, Gholamreza Salimi-Khorshidi (FMRIB, Oxford)

IT/informatics Duncan Mortimer, David Flitney, Matthew Webster, Paul McCarthy (FMRIB, Oxford), Alan Young, Jonathan Price, John Miller (CTSU, Oxford)

UK Biobank Imaging Working Group Chair: Paul Matthews (Imperial), Jimmy Bell (Westminster), Andrew Blamire (Newcastle), Rory Collins (Oxford/UK Biobank), Steve Garratt (UK Biobank), Tony Goldstone (Imperial), Nicholas Harvey (Southampton), Paul Leeson (Oxford), Karla Miller (Oxford), Stefan Neubauer (Oxford), Tim Peakman (UK Biobank), Steffen Petersen (Queen Mary College), Stephen Smith (Oxford), Cathie Sudlow (Edinburgh/UK Biobank)

Funding: MRC, Wellcome Trust, British Heart Foundation, ERC
We are also extremely grateful to all UK Biobank study participants

