ARAMIS LAB
BRAIN DATA SCIENCE

Paris Brain Institute
ICM

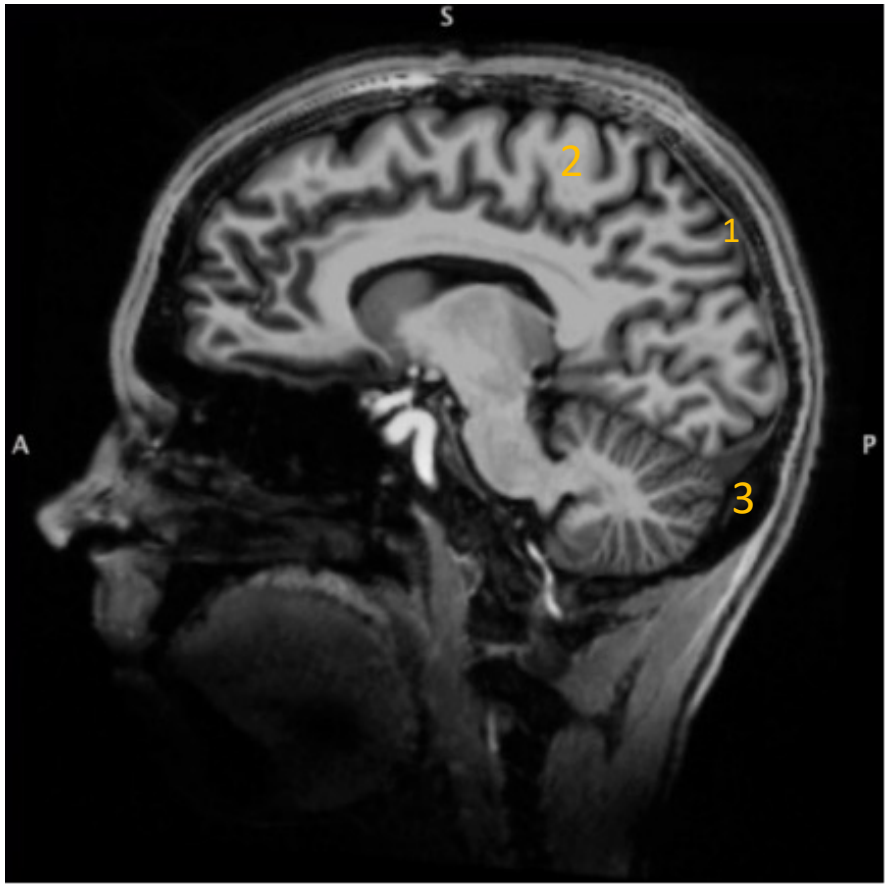
Using UK Biobank Imaging data : Focus on T1w images

February 6th, 2024

Elise Delzant, 2nd -year PhD student

Supervisor : Baptiste Couvy-Duchesne, PhD

What is a T1w image ?

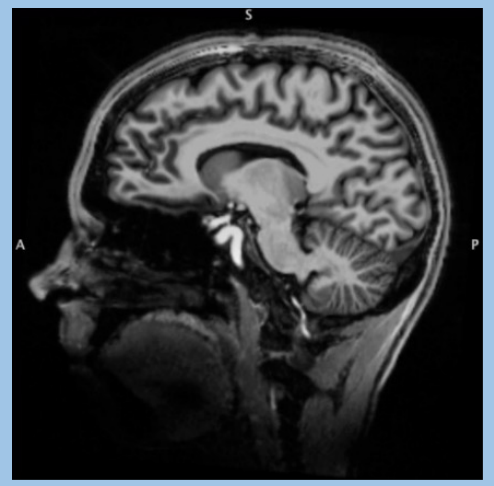


1. *Grey-matter in dark grey*
2. *White-matter in light grey*
3. *Cerebro-spinal fluid (CSF) in dark*

- Structural MRI
- Gives details of the brain :
  - enhances the signal of fatty tissue
  - suppresses signal of water
- Best depicts anatomy and can accentuate pathology

Different images for different applications

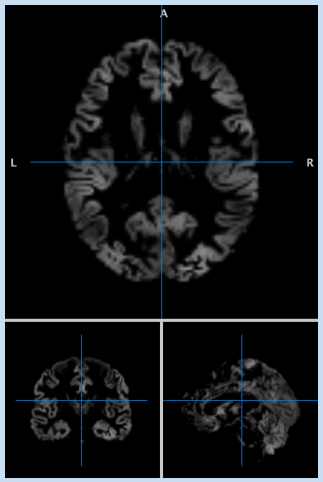
Brain MRI raw image (T1)



Unstructured data

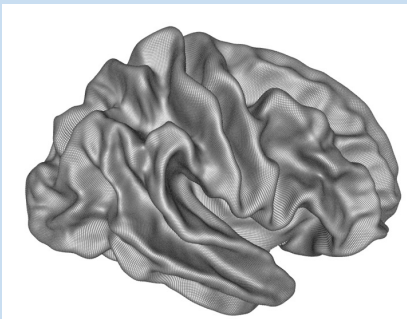
- *Deep-learning*

High dimensional processing



Volume-based

OR



Surface-based

Structured data – high dimensional

- *High performance cluster*
- *Complex statistics, machine learning*

Region of Interest

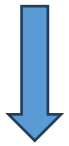
Desikian-Killiany atlas
(Desikian et al., 2006). 68 ROI

Structured data – low dimensional

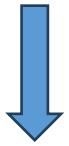
- *Simple statistics*

UK Biobank T1w images

0 Initial assessment visit (2006-2010) at which participants were recruited and consent given
1 First repeat assessment visit (2012-13)
2 Imaging visit (2014+)
3 First repeat imaging visit (2019+)



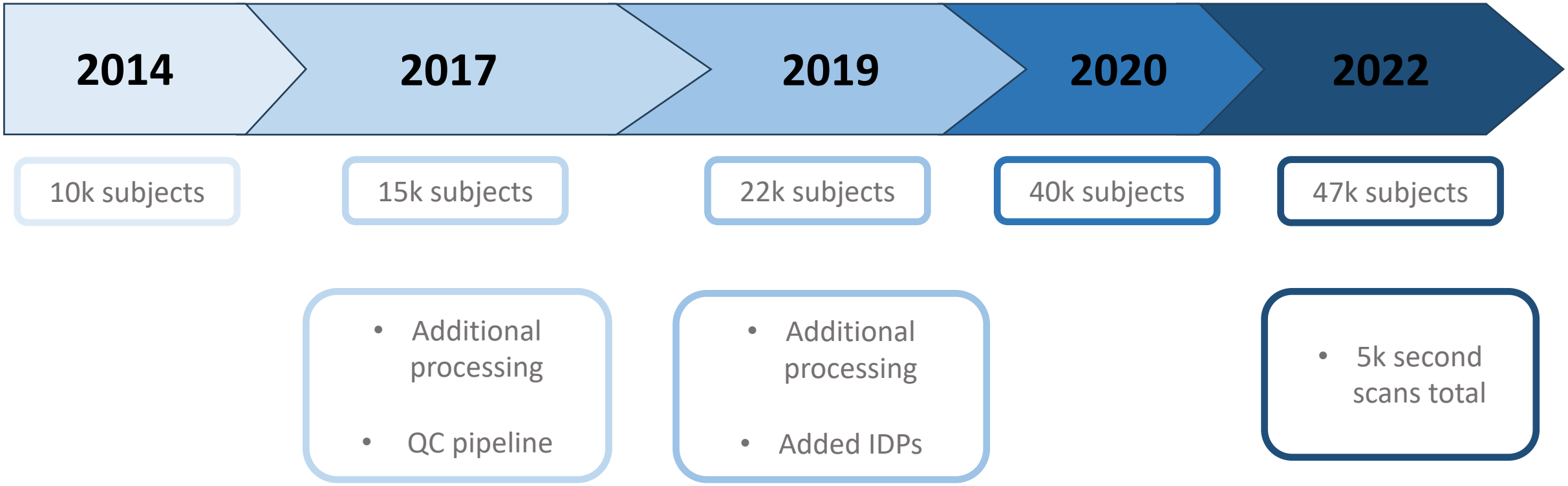
2 Imaging visit (2014+)
3 First repeat imaging visit (2019+)



5 centers :

- assessment around 4-5 hours
- provides detailed MRI measures of the brain

Data imaging release



UK Biobank T1w images

Field ID	Description	Category	
20252	T1 structural brain images - NIFTI	T1 structural brain MRI	‡

Raw + volume-based processed images

2nd visit : 66,952 participants
3rd visit : 5211 participants

2 types of T1w available :

Surface-based processed images

Field ID	Description	Category	
20263	T1 surface model files and additional structural segmentations	T1 structural brain MRI	‡

2nd visit : 43,139 participants
3rd visit : 4629 participants

Type : Bulk

biobank uk

Index | Browse | Search | Catalogues | Downloads | Login | Help

Data-Field 20252

Description: T1 structural brain images - NIFTI

Category: Assessment centre □ Imaging □ Brain MRI □ T1 structural brain MRI
Brain MRI scans + XNAT image package + Brain MRI

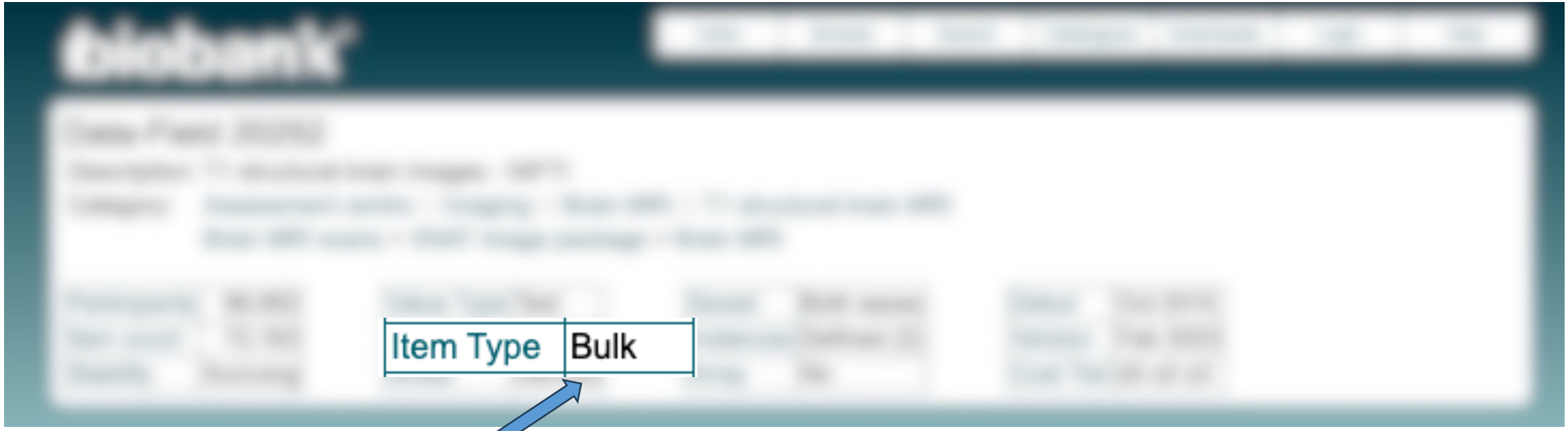
Participants	66,952
Item count	72,163
Stability	Accruing

Value Type	Text
Item Type	Bulk
Strata	Derived

Sexed	Both sexes
Instances	Defined (2)
Array	No

Debut	Oct 2015
Version	Feb 2023
Cost Tier	d3 o2 s3

Type : Bulk



- Large and complex items, such as genomic data, imaging data.

Field ID	Description	Category
20252	T1 structural brain images - NIFTI	T1 structural brain MRI ‡

- 1 fields marked ‡ are blob/bulk.

Download bulk files


```
ukbconv ukb673035.enc_ukb bulk -s20252
```

File ukb673035.bulk

Participant ID	Field_Instance_Array
3422567	20252_2_0
5321753	20252_2_0
2457842	20252_3_0

Download bulk files


```
ukbconv ukb673035.enc_ukb bulk -s20252
```

File ukb673035.bulk

Participant ID	Field_Instance_Array
3422567	20252_2_0
5321753	20252_2_0
2457842	20252_3_0

```
ukbfetch -bukb673035.bulk -ak53185r673035.key
```

Download every file from Field 20252

Download bulk files


```
ukbconv ukb673035.enc_ukb bulk -s20252
```

File ukb673035.bulk

Participant ID	Field_Instance_Array
3422567	20252_2_0
5321753	20252_2_0
2457842	20252_3_0

```
ukbfetch -bukb673035.bulk -ak53185r673035.key
```

Download every file from Field 20252

.bulk files can be edited

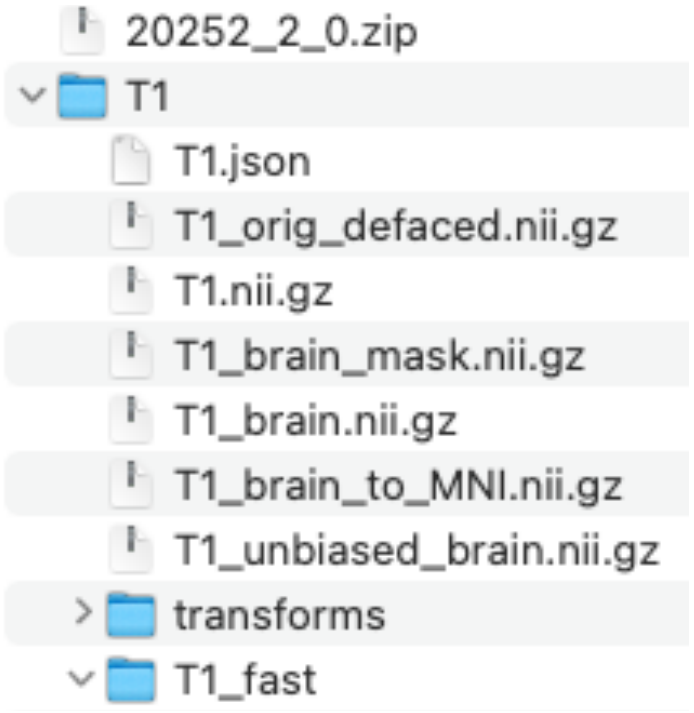
Number of files to download can be specified

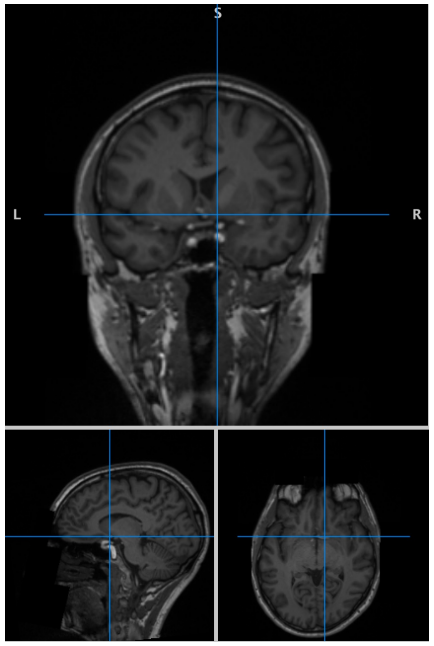
A limit of 1,000 files per ukbfetch call

```
ukbfetch -bukb673035.bulk -s1 -m1000
ukbfetch -bukb673035.bulk -s1001 -m2000
ukbfetch -bukb673035.bulk -s2001 -m3000
```

Raw + Volume-based processed images - Data Field 20252

https://biobank.ndph.ox.ac.uk/ukb/field.cgi?id=20252



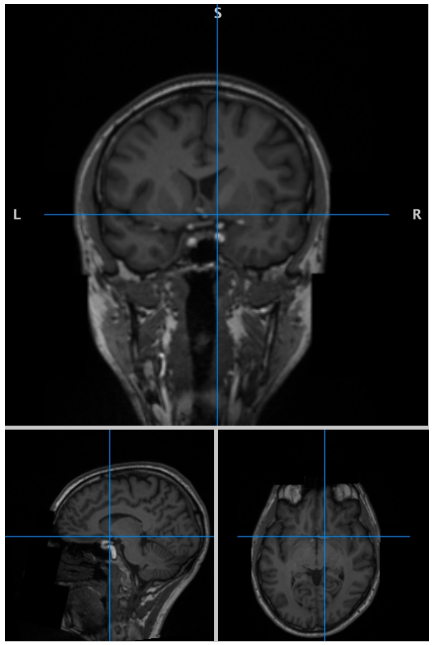


T1_orig_defaced.nii.gz: full raw T1 image after defacing

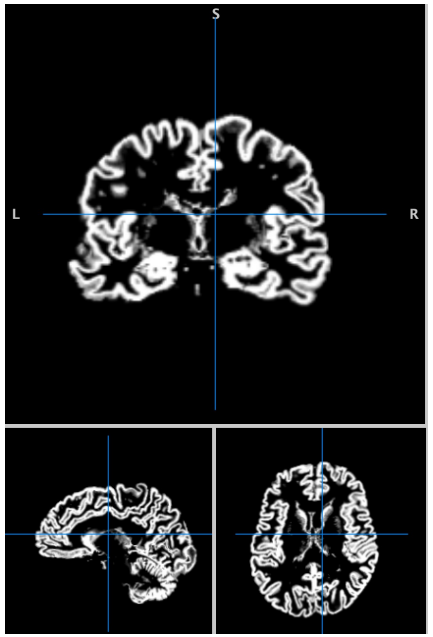
Raw + Volume-based processed images - Data Field 20252

https://biobank.ndph.ox.ac.uk/ukb/field.cgi?id=20252

- 20252_2_0.zip
- T1
  - T1.json
  - T1_orig_defaced.nii.gz
  - T1.nii.gz
  - T1_brain_mask.nii.gz
  - T1_brain.nii.gz
  - T1_brain_to_MNI.nii.gz
  - T1_unbiased_brain.nii.gz
  - transforms
  - T1_fast
    - T1_brain_seg.nii.gz
    - T1_brain_pve_0.nii.gz
    - T1_brain_pve_1.nii.gz
    - T1_brain_pve_2.nii.gz
    - T1_brain_bias.nii.gz



T1_orig_defaced.nii.gz: full raw T1 image after defacing

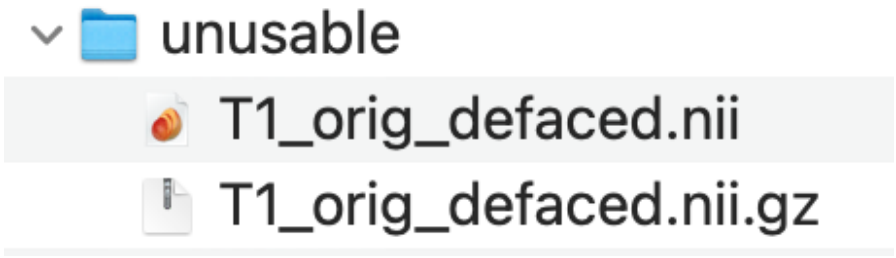


T1_brain_pve_1.nii.gz: segmentation of grey-matter

Jenkinson et al., 2012

50 MB per participant

Semi-automated QC



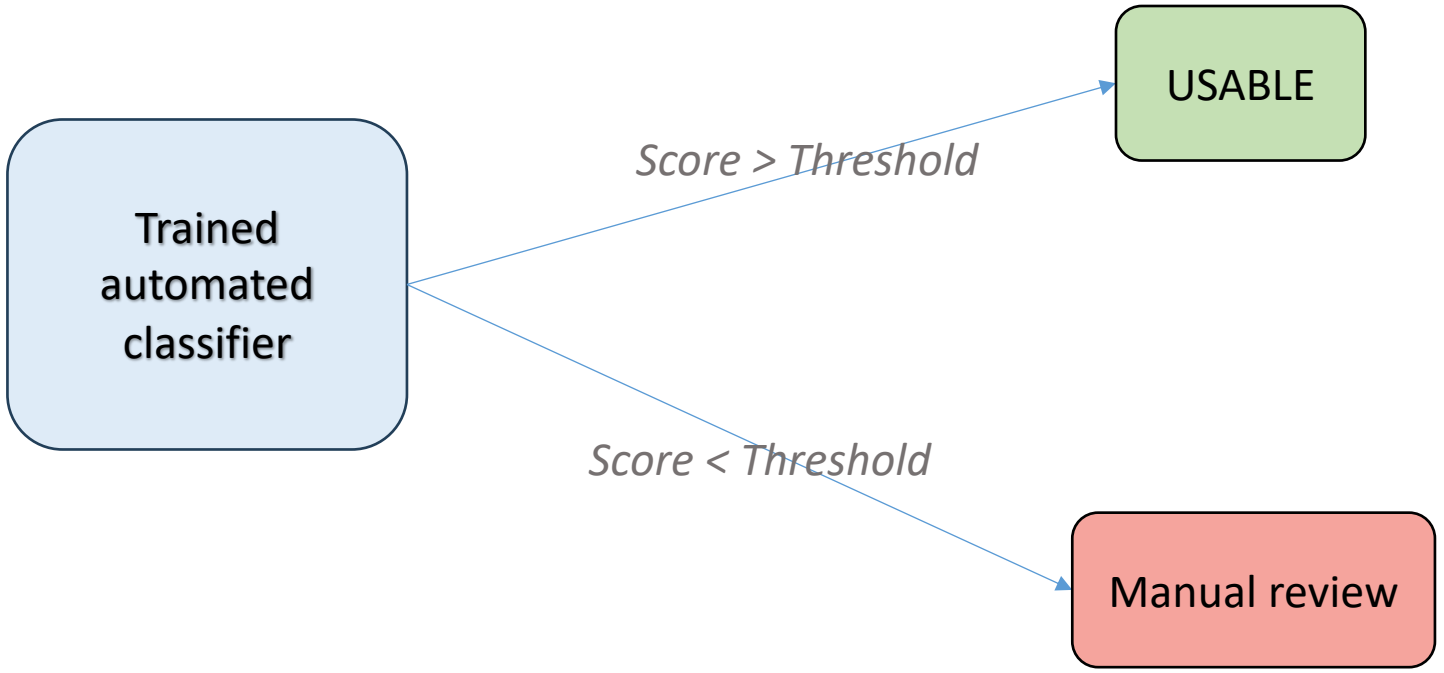
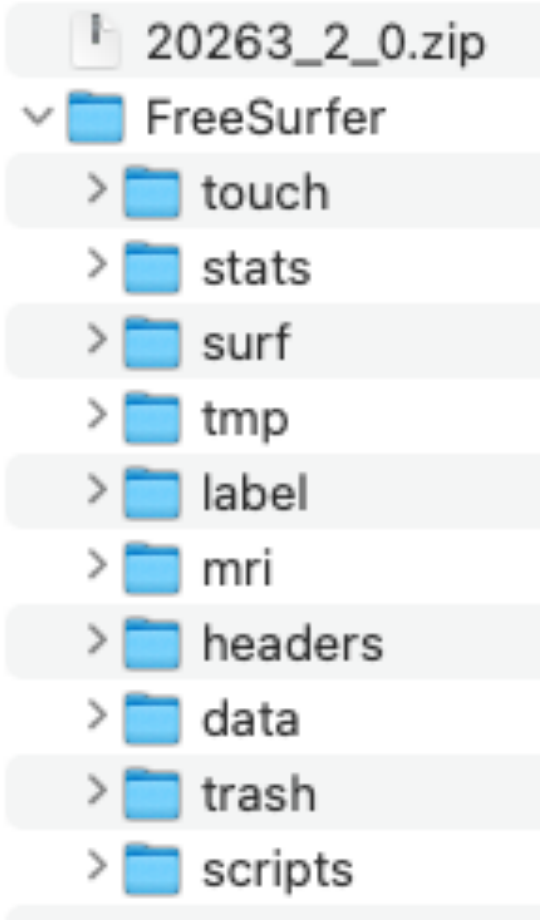


Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank, Alfaro-Almagro et al. , NeuroImage, 2018

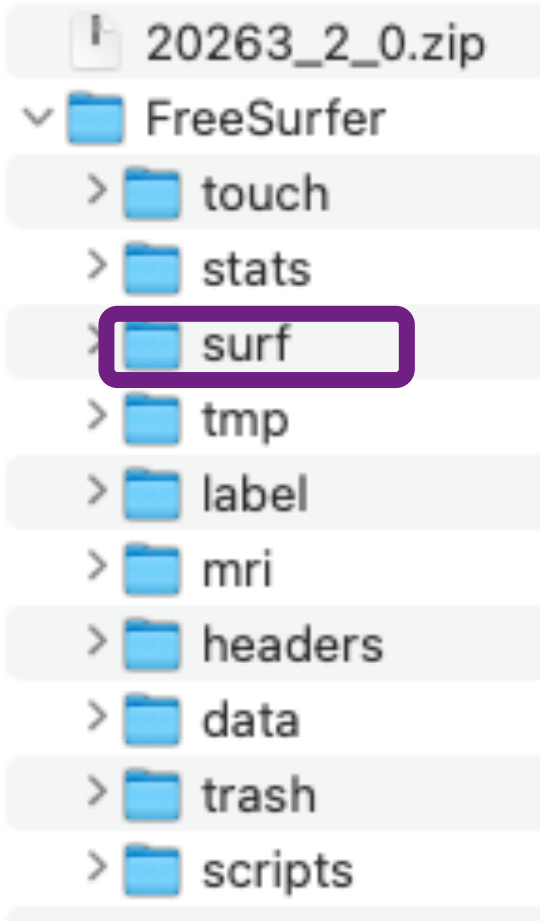
Surface-based processed images-Data Field 20263

https://biobank.ndph.ox.ac.uk/ukb/field.cgi?id=20263



Surface-based processed images-Data Field 20263

https://biobank.ndph.ox.ac.uk/ukb/field.cgi?id=20263

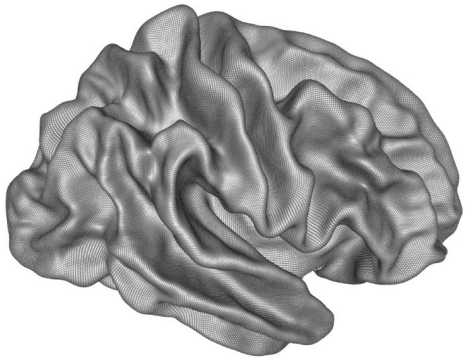


FreeSurfer

~ 6h per subject

Surface model files :

- Cortical surface : left and right hemisphere area, volume, cortical thickness
- Subcortical data



350 MB per participant

Imaging-derived phenotypes (IDP)

Derived measures of brain structure

Global Measures :

Total Brain Volume

Grey-matter Volume

White-matter Volume

Regional measures :

Left Hippocampus Volume

Cortical Thickness

Goal :

Using imaging data in non-imaging analysis

- *Association of a wide range of individual chronic diseases and their multimorbidity with brain volumes in the UK Biobank: A cross-sectional study, X. Shang et al., eClinicalMedecine, 2022*
- *Structural brain imaging correlates of general intelligence in UK Biobank, Cox et al., Intelligence, 2019*

ARAMIS LAB
BRAIN DATA SCIENCE

Example of an application with T1w images

Objectives

Which of FSL or FreeSurfer capture the most information ?

FreeSurfer

vs.

N=42,272

X 350 MB = **15 TeraBytes**

N=42,272

X 50 MB = **2 TeraBytes**

Objectives

Which of FSL or FreeSurfer capture the most information ?

FreeSurfer

vs.

N=42,272

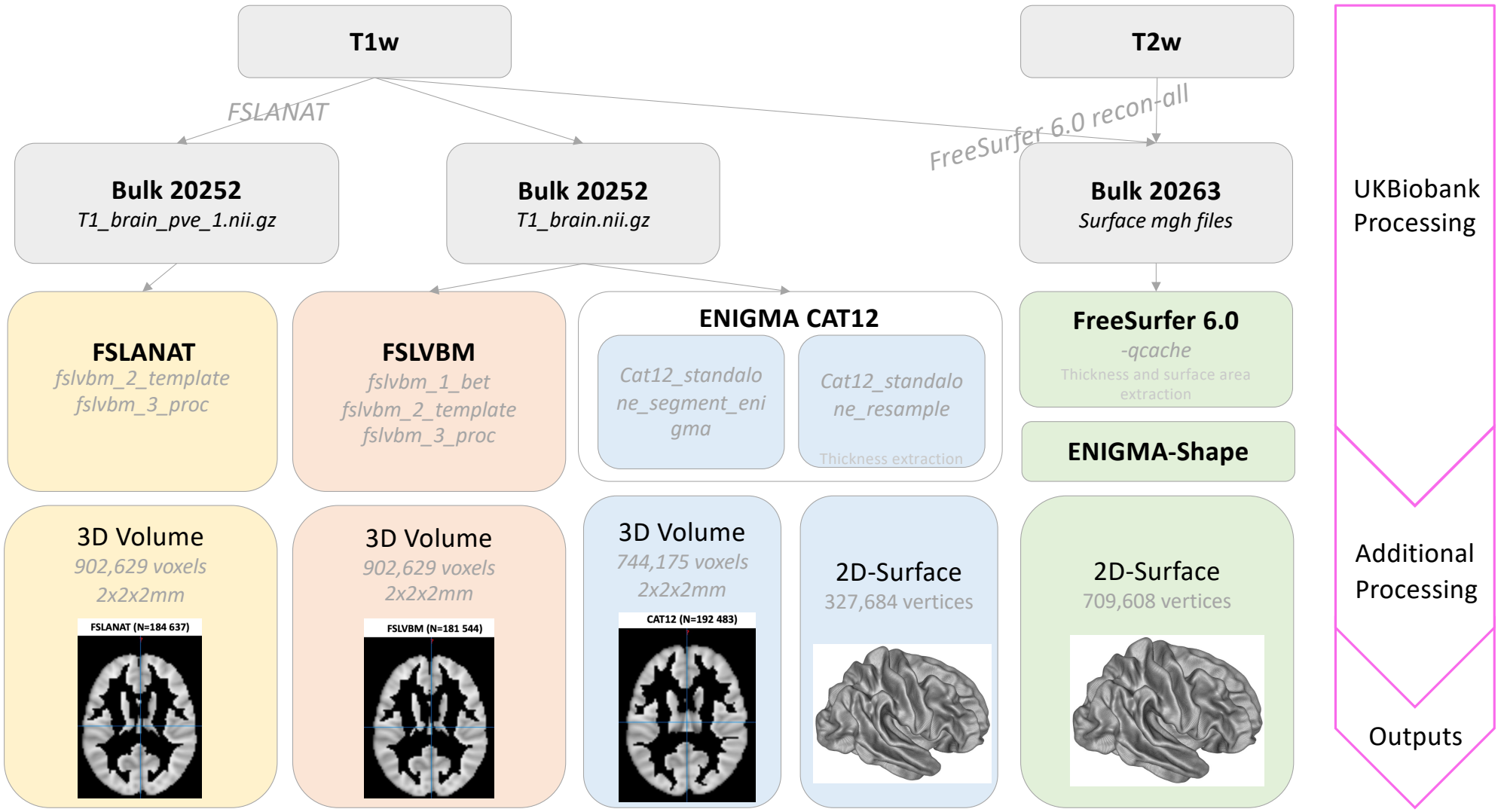
X 350 MB = **15 TeraBytes**

N=42,272

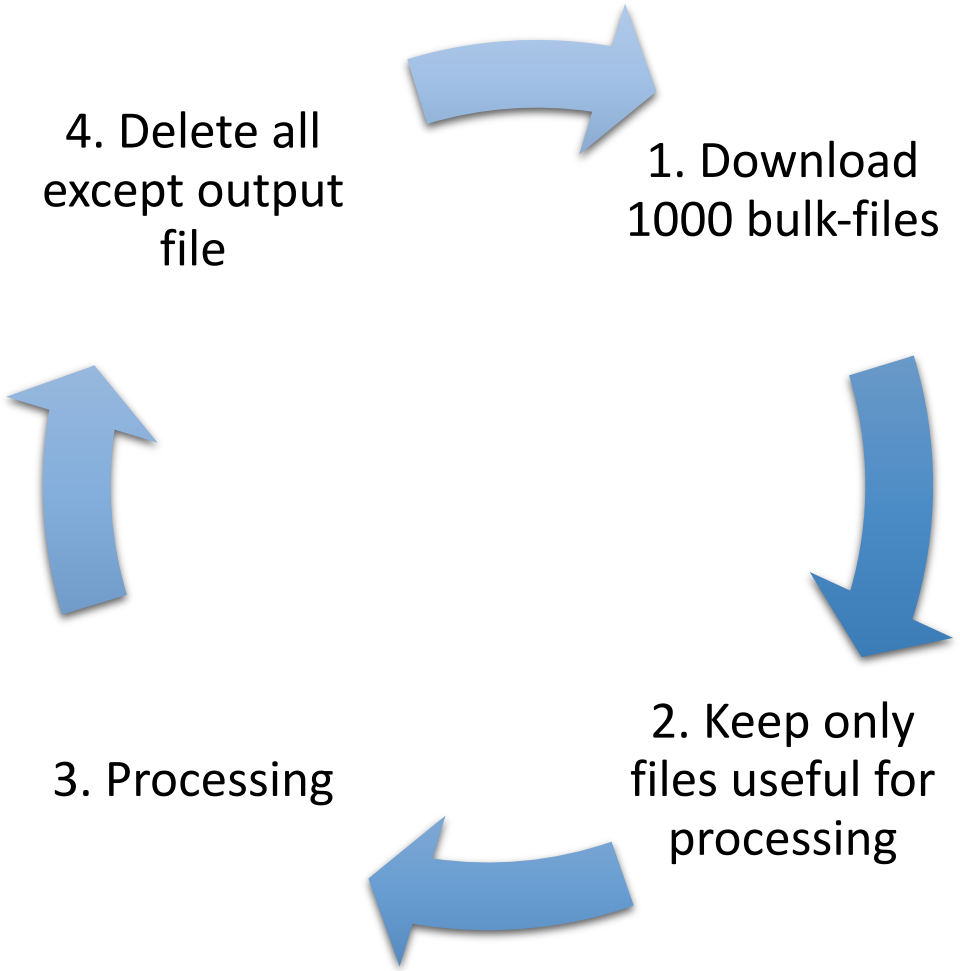
X 50 MB = **2 TeraBytes**

\+

Detailed processing



Data management



FreeSurfer vs. FSL

N=42,272

X 350 MB = **15 TeraBytes**

N=42,272

X 50 MB = **2 TeraBytes**

1 file per subject per processing

FSL : 143G Total

Comparing processing

We estimated, for each processing, the percentage of variance of a trait captured by all brain features

Potential confounders

- *Signal to noise ratio*
- *Time difference to first MRI*
- *Discrepancy between T1 and template*
- *Mean head motion*
- *X and Y brain position*

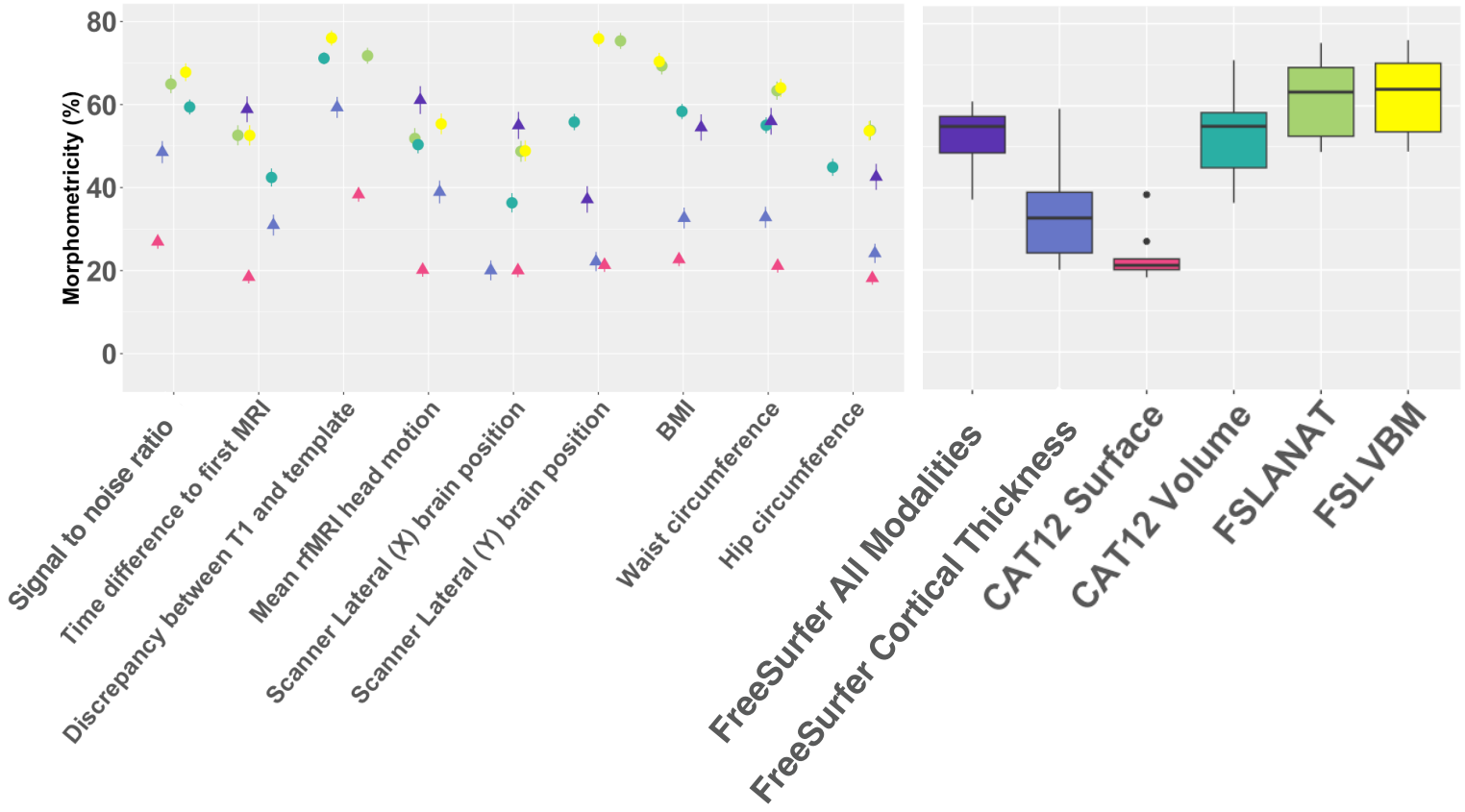
Body-size covariates

- *BMI*
- *Waist circumference*
- *Hip circumference*

Traits of interest

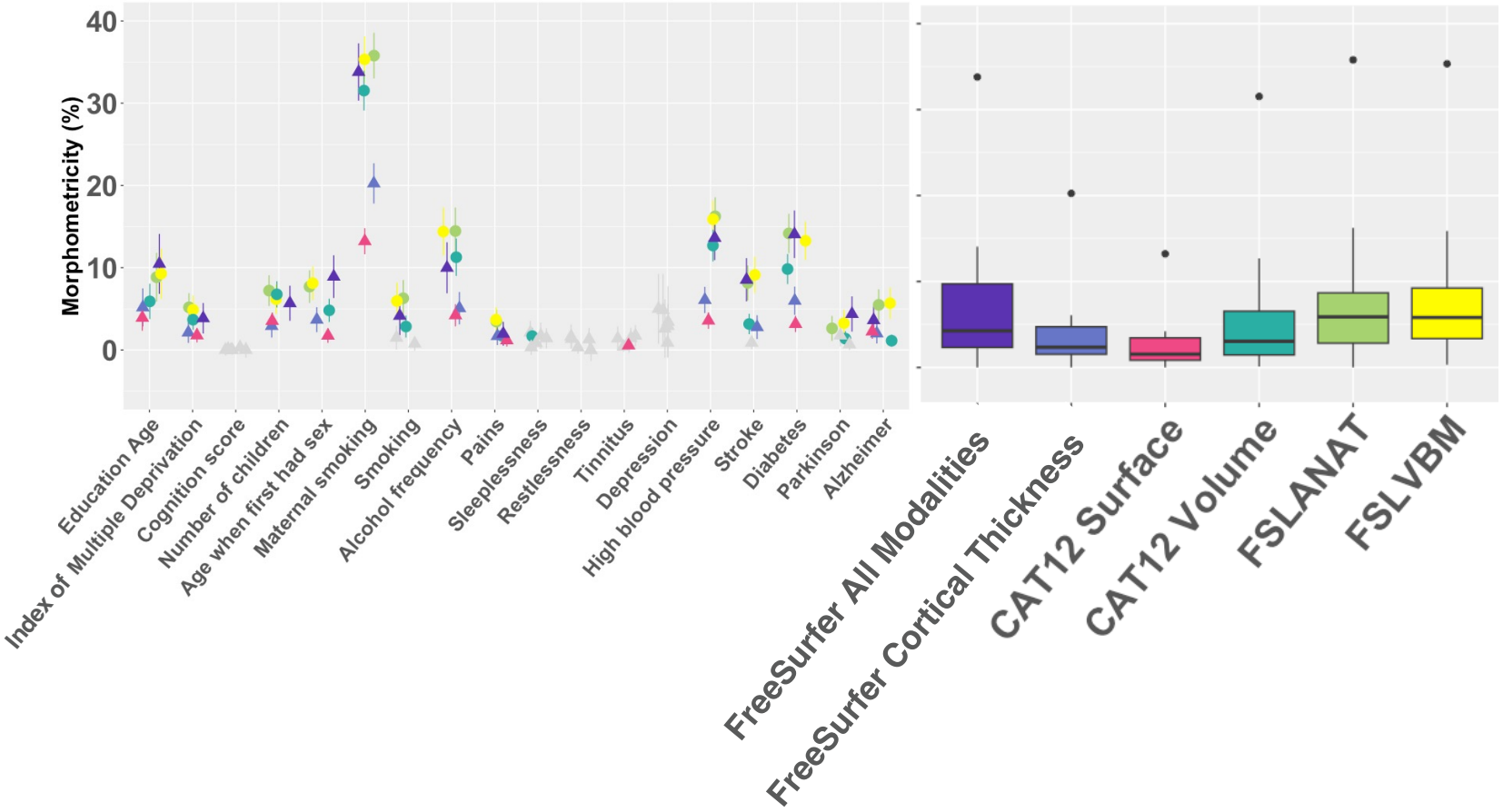
- *Education Age*
- *Index of deprivation*
- *Cognition score*
- *Number of children*
- *Age First had sex*
- *Maternal smoking*
- *Smoking*
- *Alcohol fequency*
- *Pains*
- *Sleeplessness*
- *Restlessness*
- *Tinnitus*
- *Depression*
- *High blood pressure*
- *Stroke*
- *Diabetes*
- *Parkinson*
- *Alzheimer*

Results : potential confounders



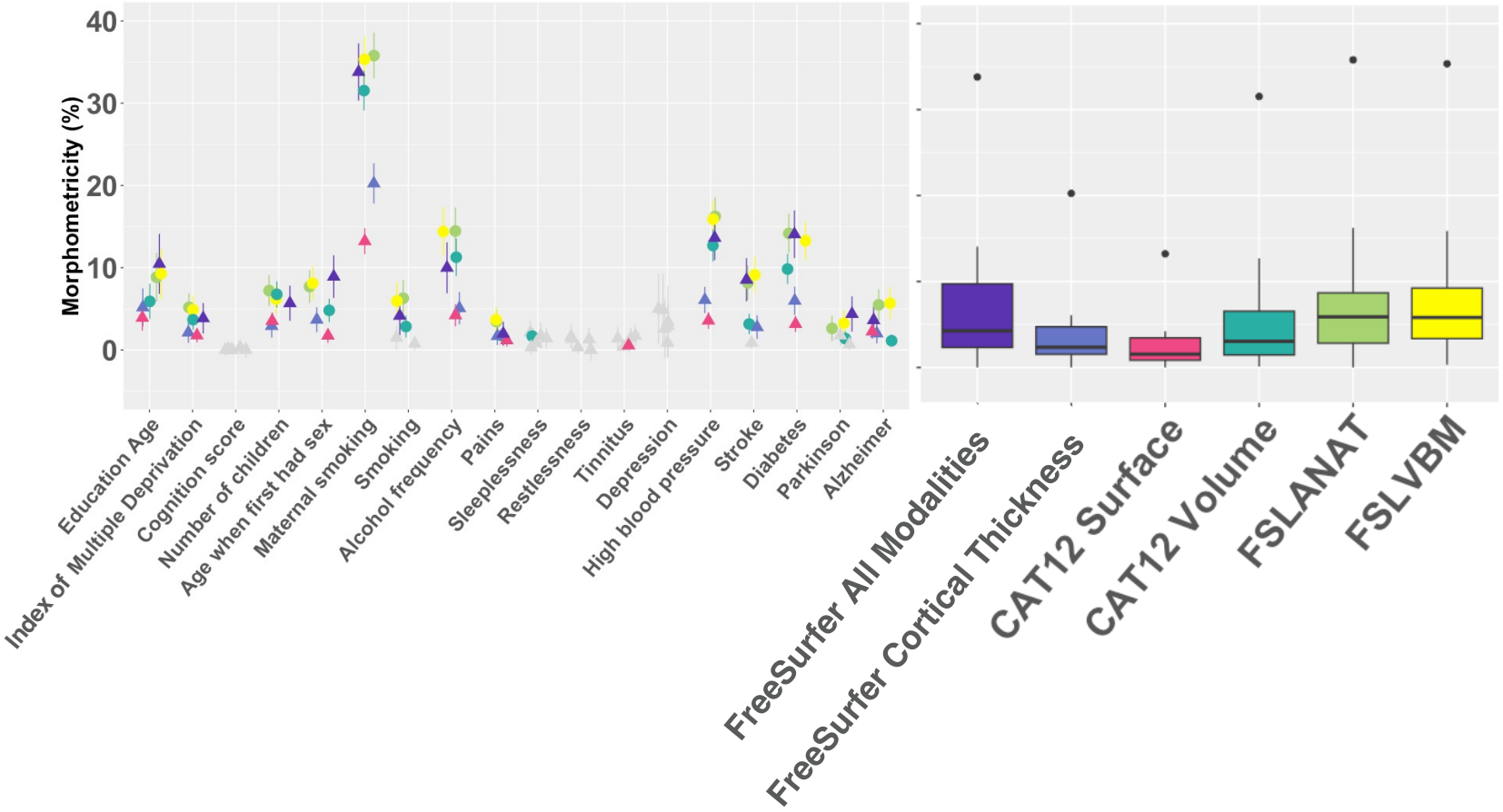
- All confounders should be included in grey-matter analysis of the UK Biobank

Results : traits of interest



- The choice of processing impacts the percentage of variance estimated

Results : traits of interest



- The choice of processing impacts the percentage of variance estimated

More resources

UK Biobank Brain Imaging Documentation

Version 1.9
September 2022

primary documentation authors:
Stephen M. Smith, Fidel Alfaro-Almagro and Karla L. Miller
Wellcome Centre for Integrative Neuroimaging (WIN-FMRIB), Oxford University on behalf of UK Biobank

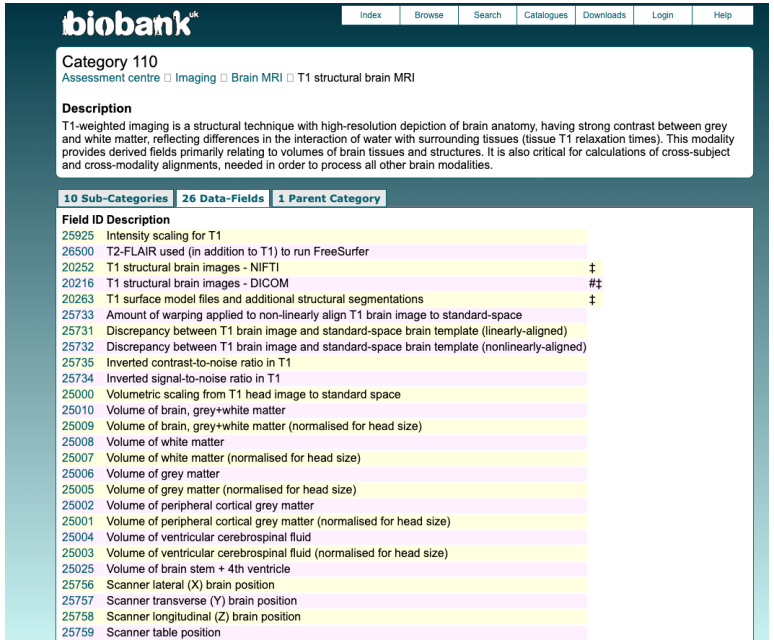
Resource | Published: 19 September 2016

Multimodal population brain imaging in the UK Biobank prospective epidemiological study

Karla L Miller ✉, Fidel Alfaro-Almagro, Neal K Bangerter, David L Thomas, Essa Yacoub, Junqian Xu, Andreas J Bartsch, Saad Jbabdi, Stamatios N Sotiropoulos, Jesper L R Andersson, Ludovica Griffanti, Gwenaëlle Douaud, Thomas W Okell, Peter Weale, Iulius Dragonu, Steve Garratt, Sarah Hudson, Rory Collins, Mark Jenkinson, Paul M Matthews & Stephen M Smith

Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank

Fidel Alfaro-Almagro,[a,*] Mark Jenkinson,[a] Neal K. Bangerter,[b] Jesper L.R. Andersson,[a] Ludovica Griffanti,[a] Gwenaëlle Douaud,[a] Stamatios N. Sotiropoulos,[a,c] Saad Jbabdi,[a] Moises Hernandez-Fernandez,[a] Emmanuel Vallee,[a] Diego Vidaurre,[d] Matthew Webster,[a] Paul McCarthy,[a] Christopher Rorden,[e] Alessandro Daducci,[f,g] Daniel C. Alexander,[h] Hui Zhang,[h] Iulius Dragonu,[i] Paul M. Matthews,[j,k] Karla L. Miller,[a] and Stephen M. Smith[a]



https://biobank.ndph.ox.ac.uk/ukb/label.cgi?id=110

Thank you for listening !

Thank you to all UK Biobank staff and participants !

Any questions?

This research has been conducted using the UK Biobank Resource under Application Number 53185