

Normative modelling: the what, when and why

Dr. Johanna Bayer

Post Doctoral Researchers
Predictive Clinical Neuroscience Lab



Predictive Clinical Neuroscience Lab

Slides by: @DrHannahSavage Dr. Charlotte Fraza @CharFraza Dr. Barbora Rehák Bučková @BarboraRehak

OUTLINE

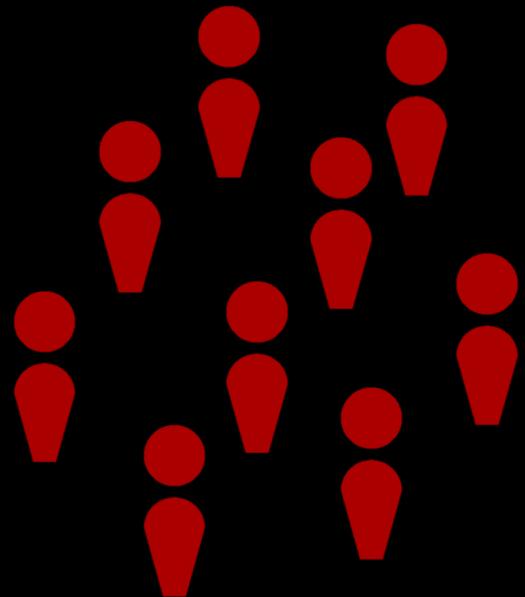
1. Short and conceptual introduction to normative modelling
2. How to train your own normative models
3. How to use our pretrained models
4. Methodological developments
5. Tutorial (Jupyter notebooks)

Resources:

- PCN Toolkit: <https://pcn toolkit.readthedocs.io/en/latest/>
- PCN Portal: <https://pcn portal.dccn.nl/>
- Tutorials and Repository for this talk:
https://github.com/likeajumpope/Tutorial_normative_modelling/blob/master/README.md
- Pre-fitted models: <https://github.com/predictive-clinical-neuroscience/braincharts>
- More tutorials: <https://github.com/predictive-clinical-neuroscience/PCN toolkit-demo>

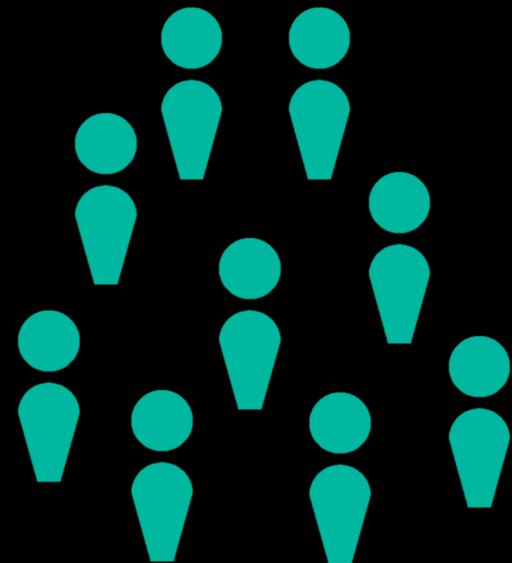
THE SHORTCOMINGS OF CLASSICAL STATISTICS

CASE



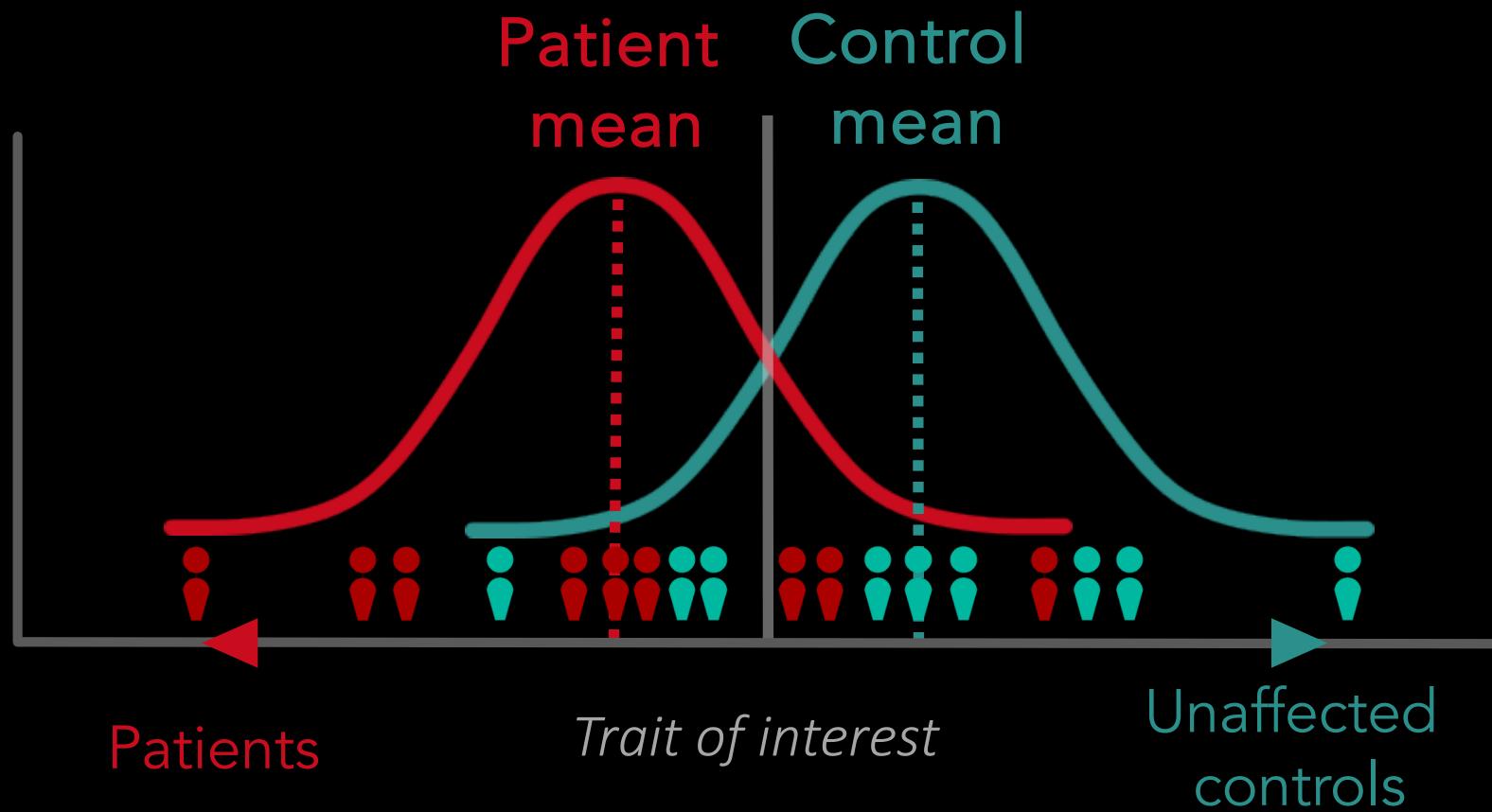
Patients

CONTROL

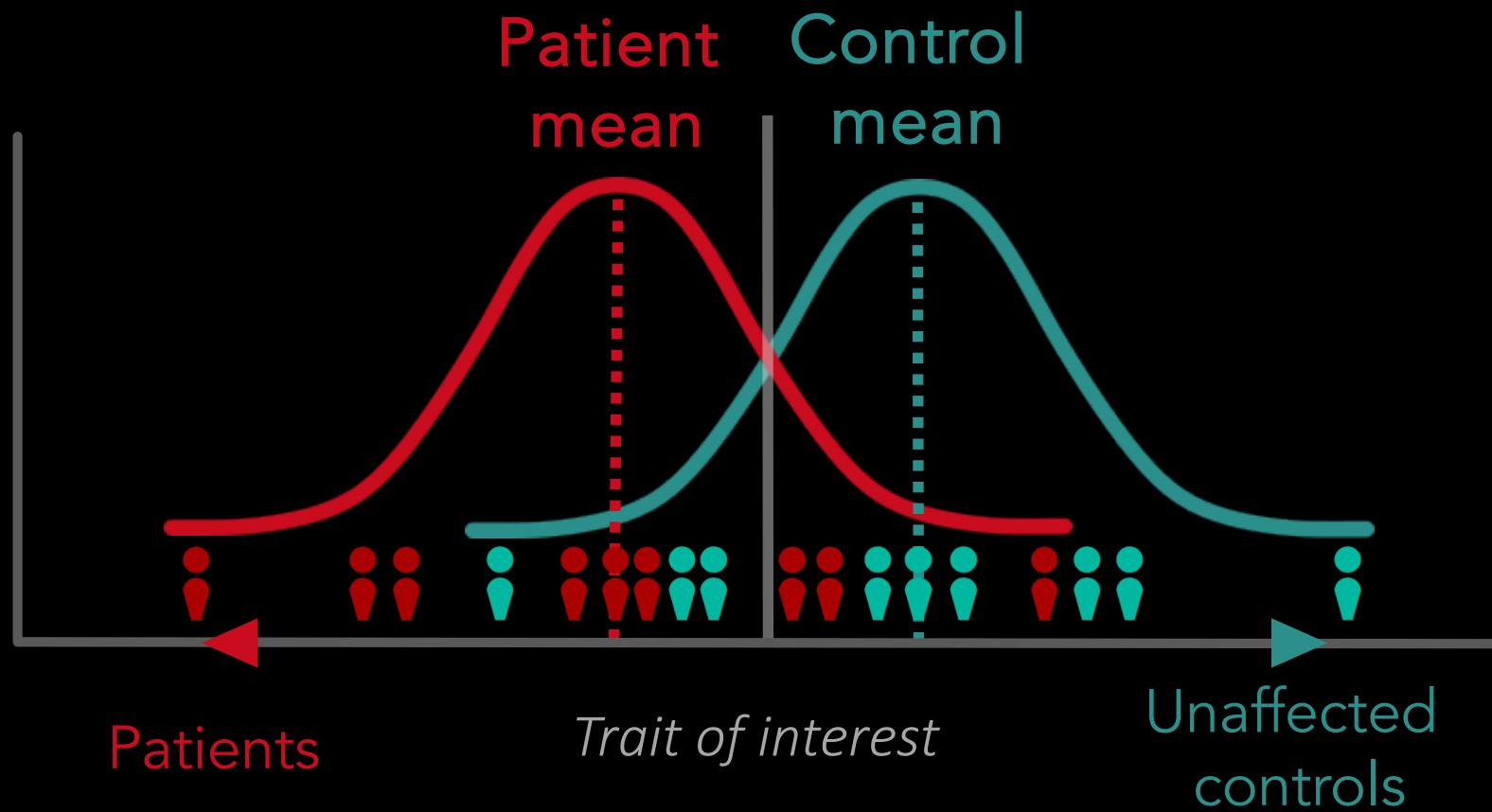


Unaffected
controls

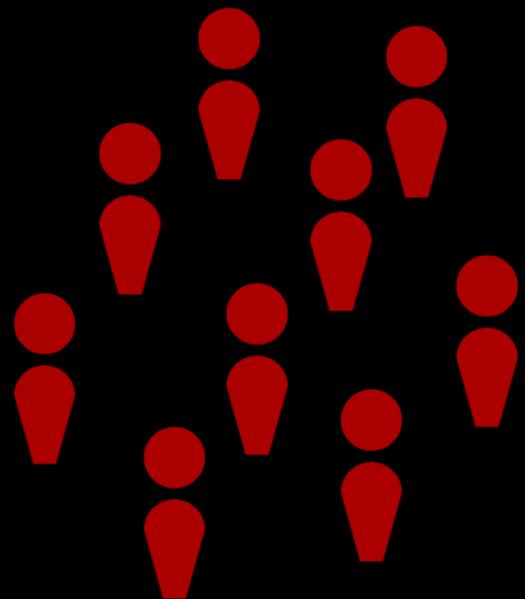
THE SHORTCOMINGS OF CLASSICAL STATISTICS



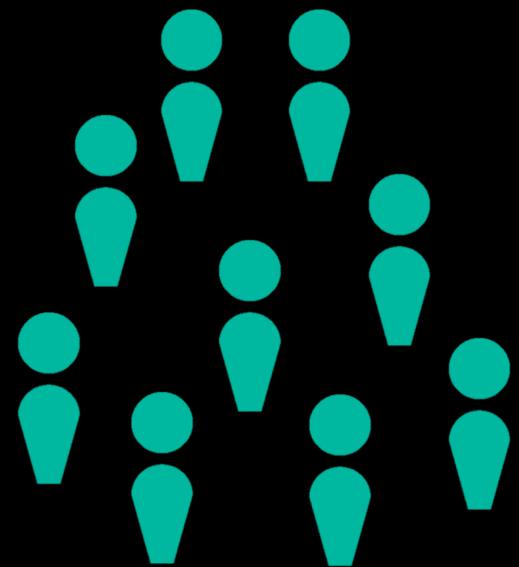
THE SHORTCOMINGS OF CLASSICAL STATISTICS



ON HETEROGENEITY



Patients

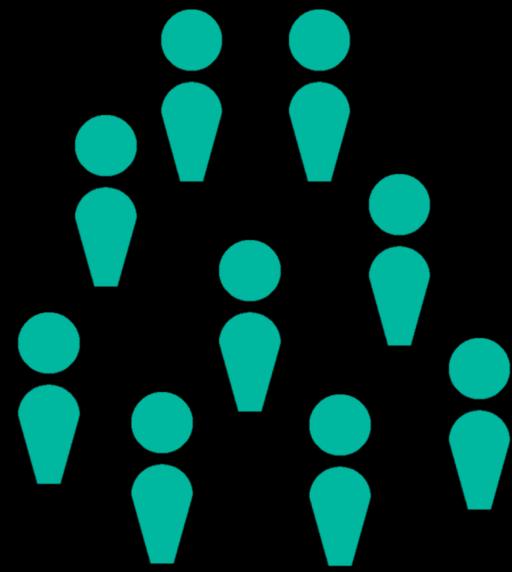


Unaffected
controls

ON HETEROGENEITY



Patients

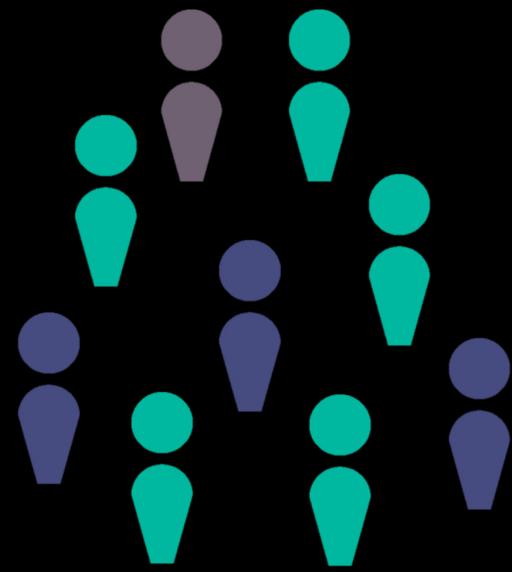


Unaffected
controls

ON HETEROGENEITY



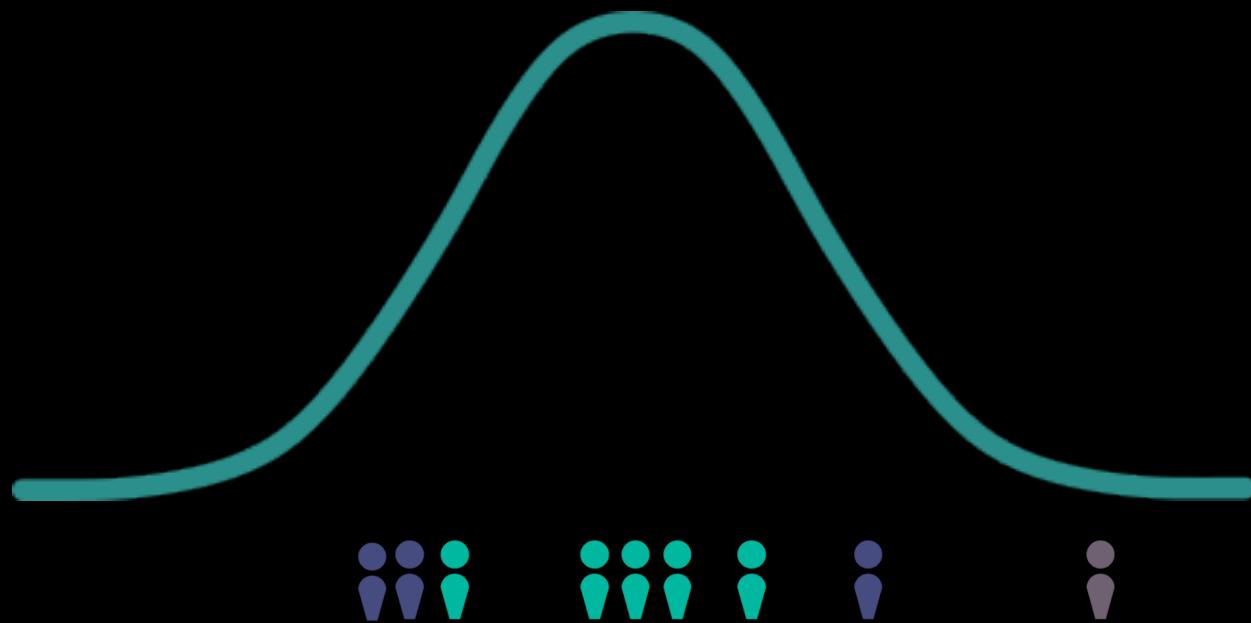
Patients



Unaffected
controls

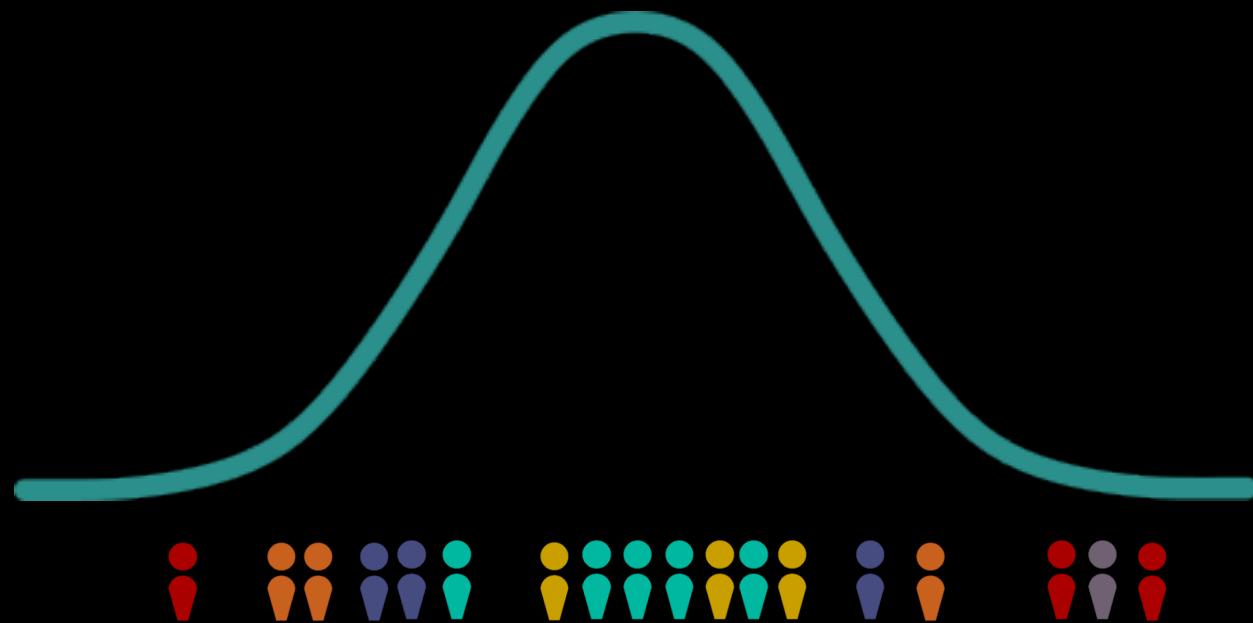
HETEROGENEITY

NORMATIVE MODELLING



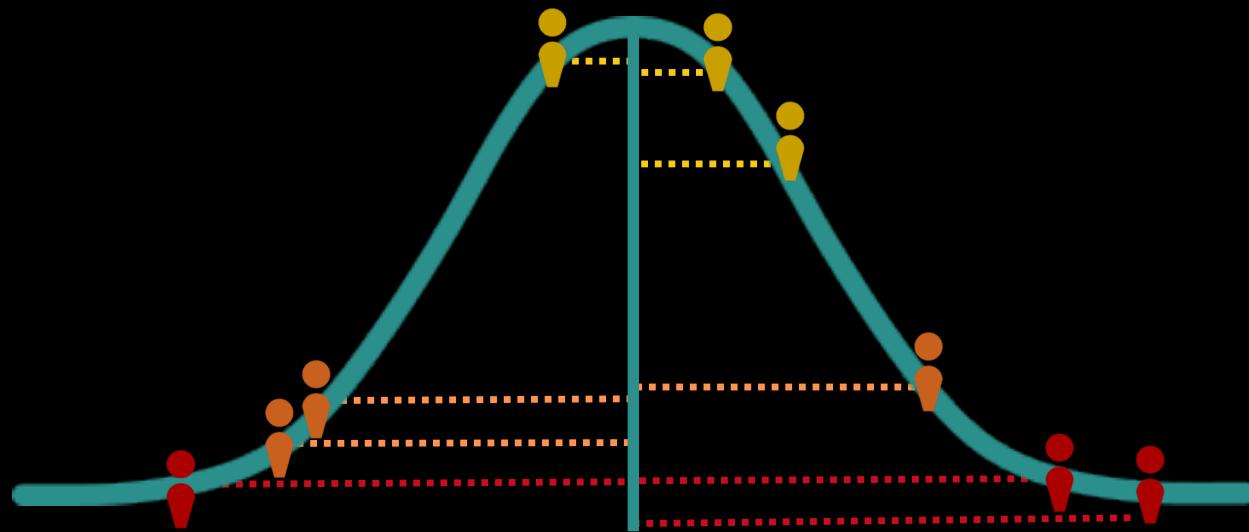
HETEROGENEITY

NORMATIVE MODELLING



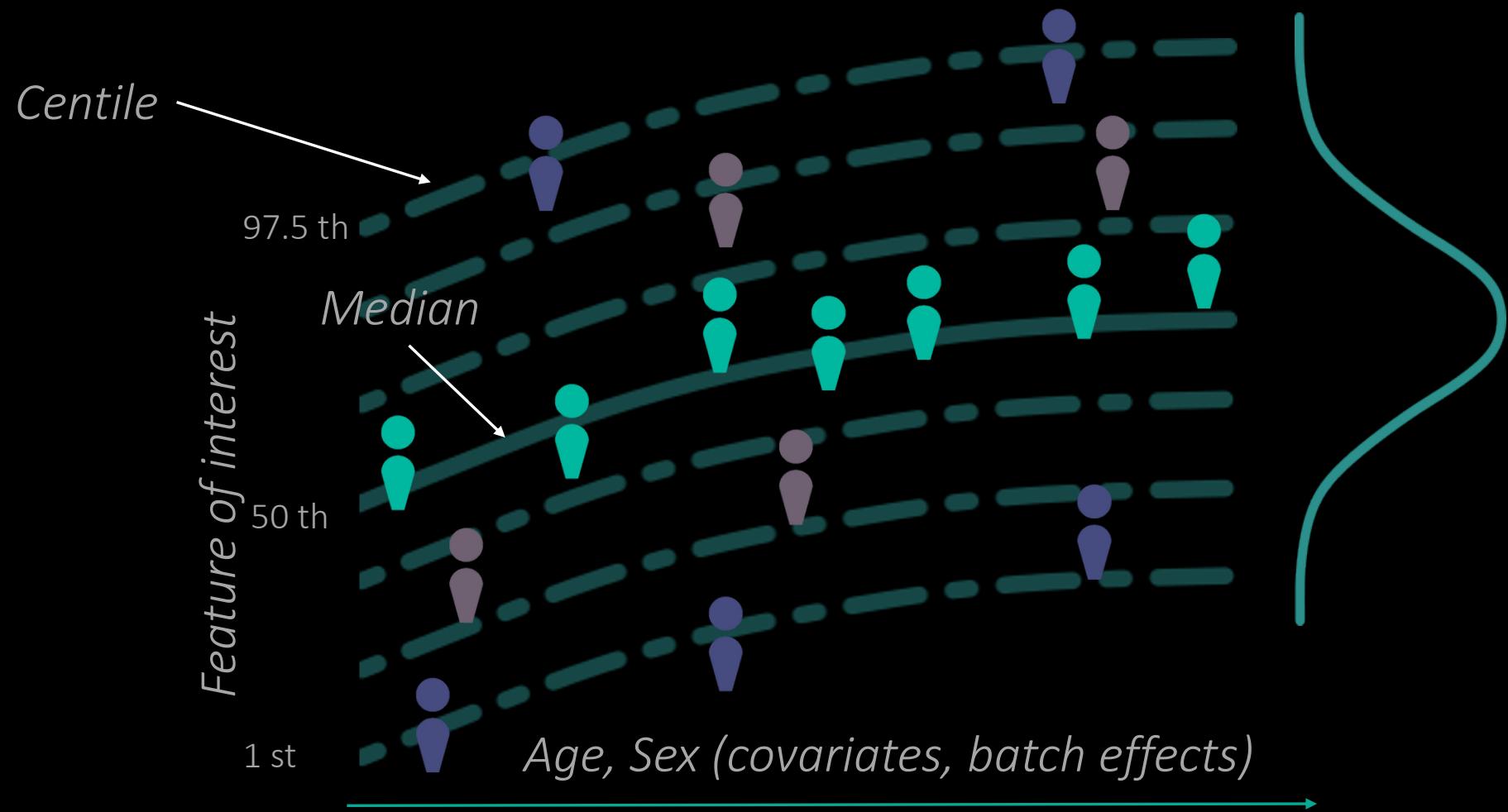
HETEROGENEITY

NORMATIVE MODELLING



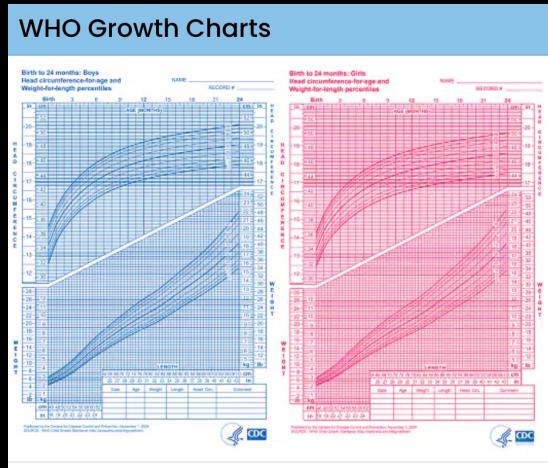
HETEROGENEITY

NORMATIVE MODELLING



HETEROGENEITY

NORMATIVE MODELLING



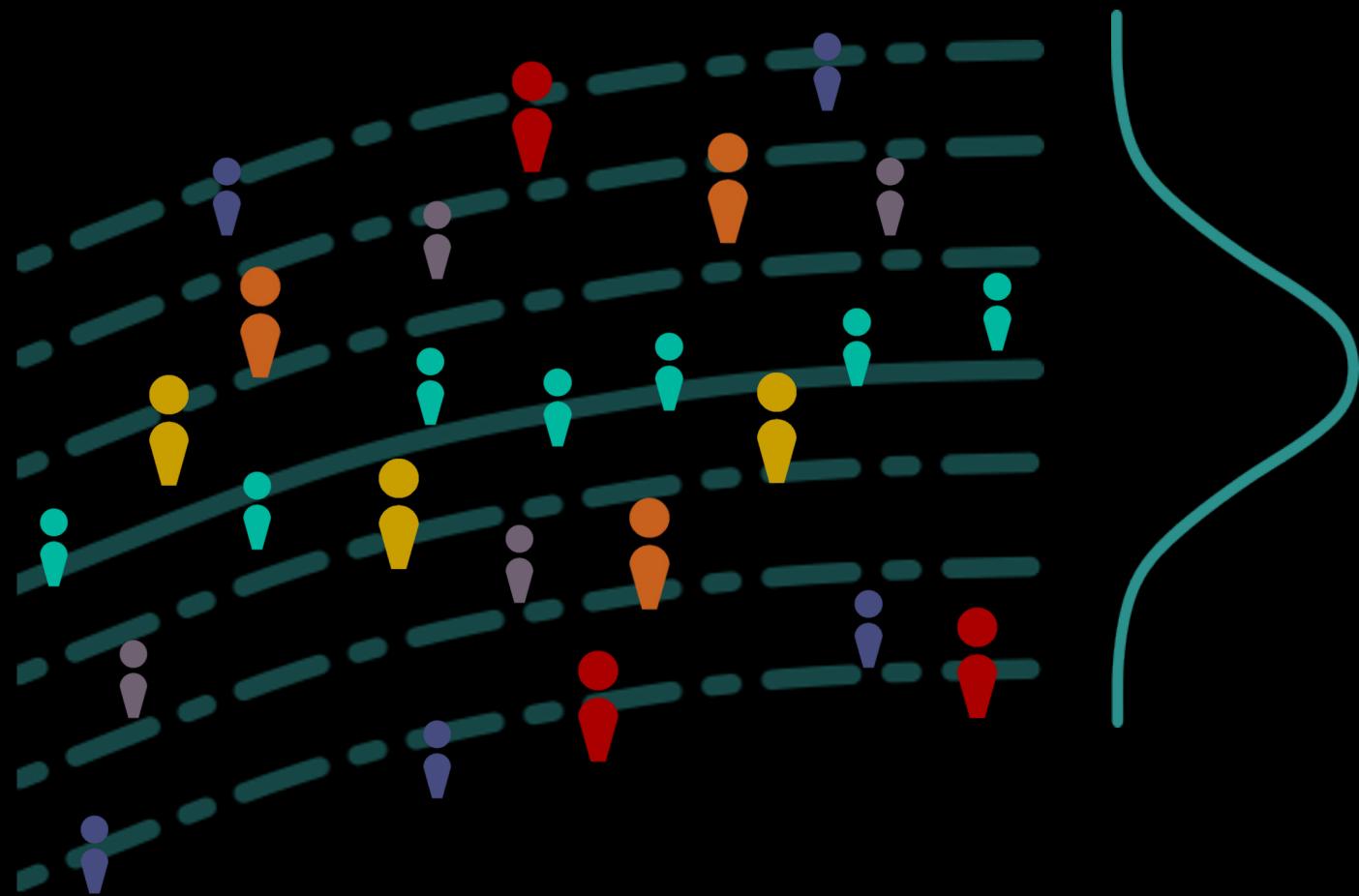
Feature of interest



HETEROGENEITY

NORMATIVE MODELLING

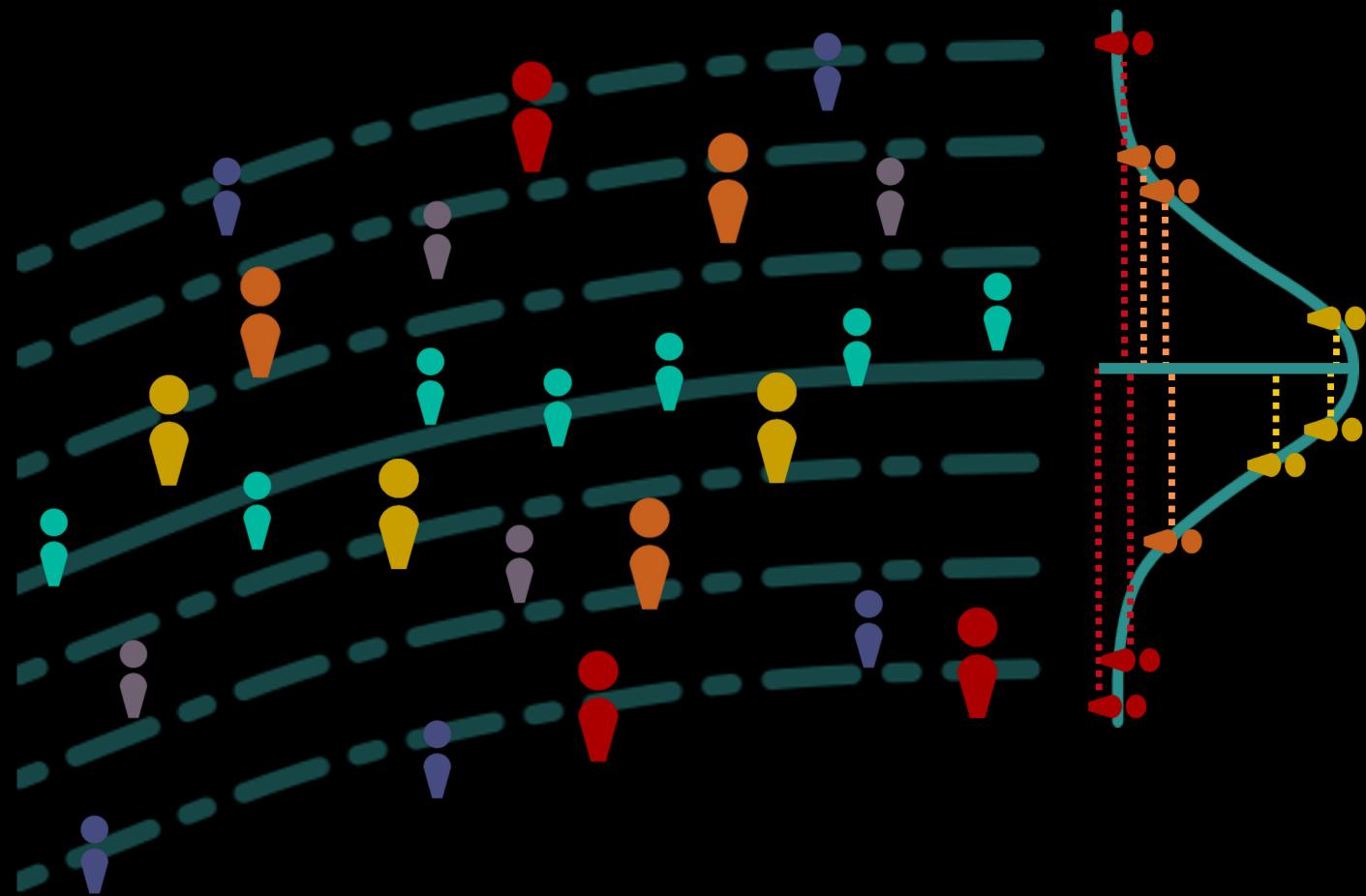
Feature of interest



HETEROGENEITY

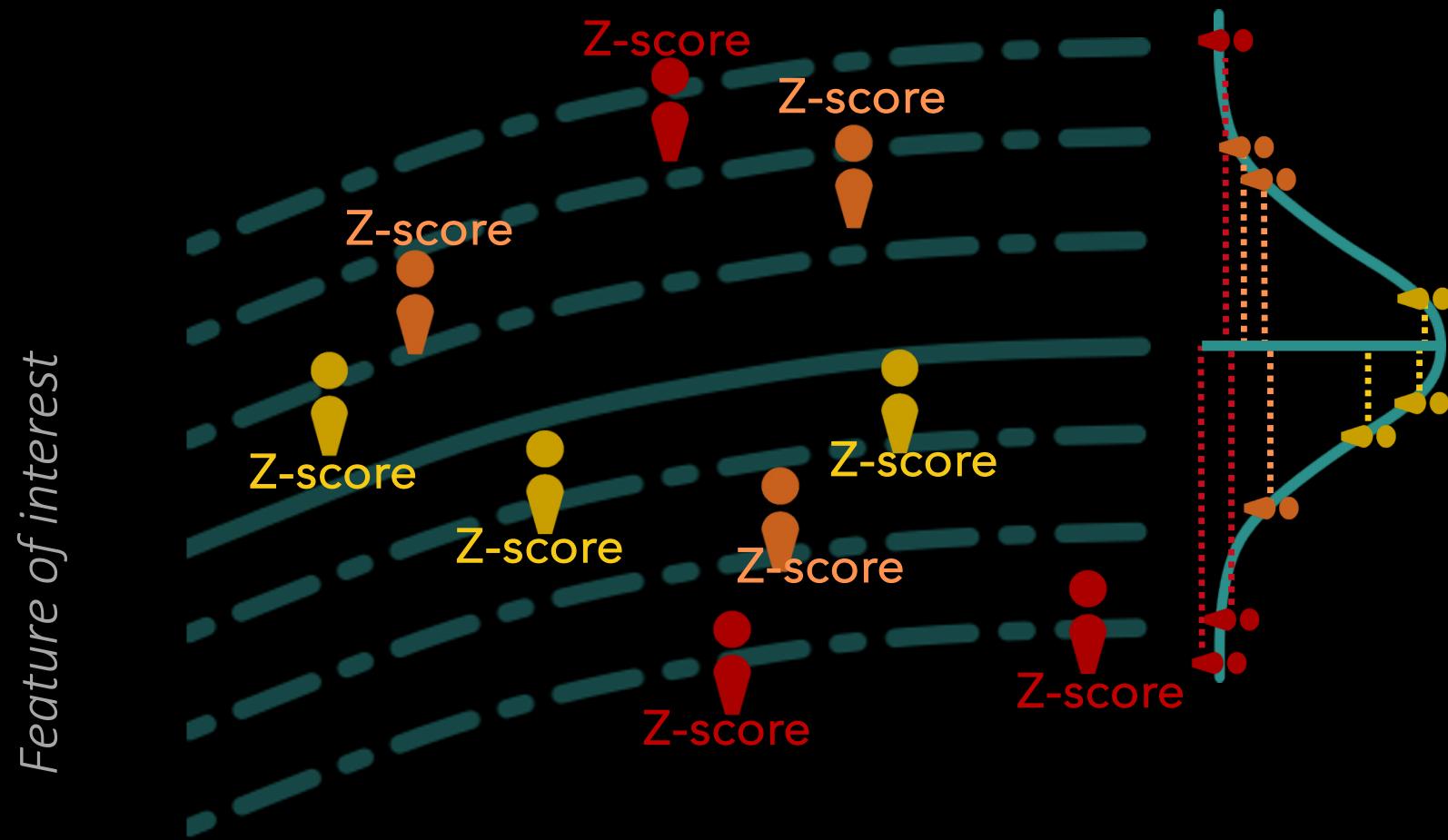
NORMATIVE MODELLING

Feature of interest



HETEROGENEITY

NORMATIVE MODELLING



HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS & DEVELOPMENT

Training your
own models

Applying pre-
trained models

Methodology

Training your
own models

Applying pre-
trained models

Methodology

HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS & DEVELOPMENT

Y

Brain
(BOLD signal in
voxel, ROI)

$$Y = f(X, \theta) + \varepsilon$$

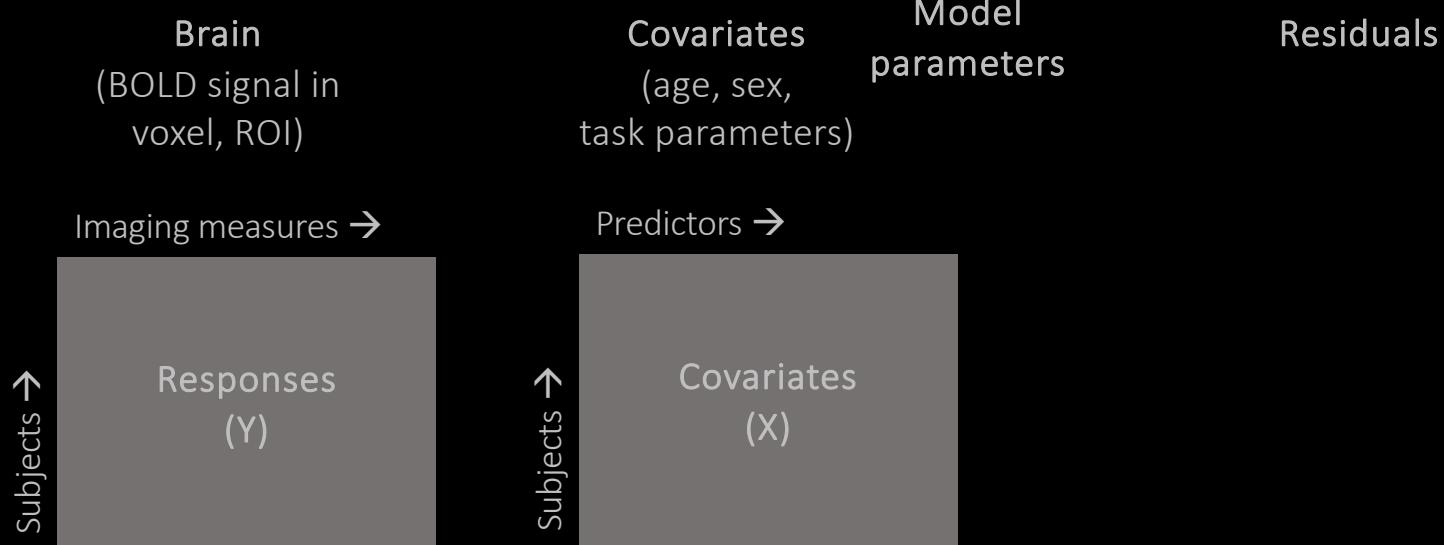
Brain
(BOLD signal in
voxel, ROI)

Covariates
(age, sex,
task parameters)

Model
parameters

Residuals

$$Y = f(X, \theta) + \epsilon$$

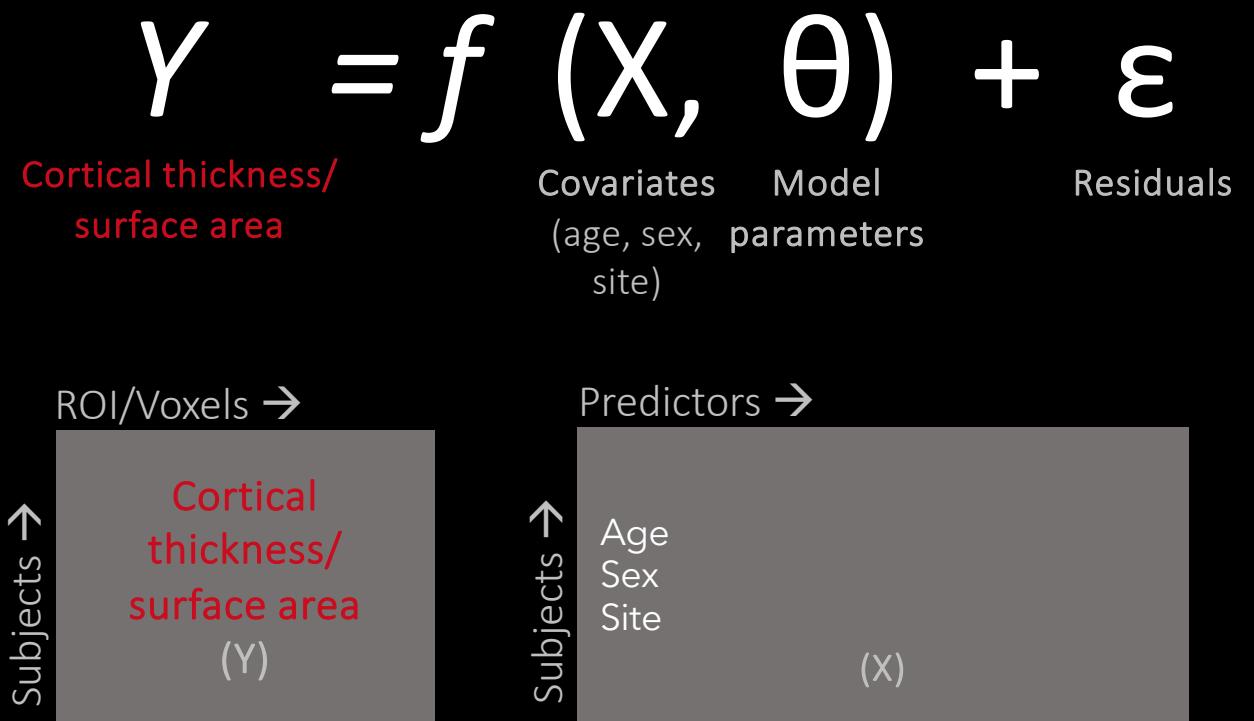


HETEROGENEITY NORMATIVE MODELLING

TRAINING YOUR OWN MODELS

APPLICATIONS & DEVELOPMENT

- Training and evaluating the normative model on local dataset



- Training and evaluating the normative model on local dataset

$$Y = f(X, \theta) + \epsilon$$

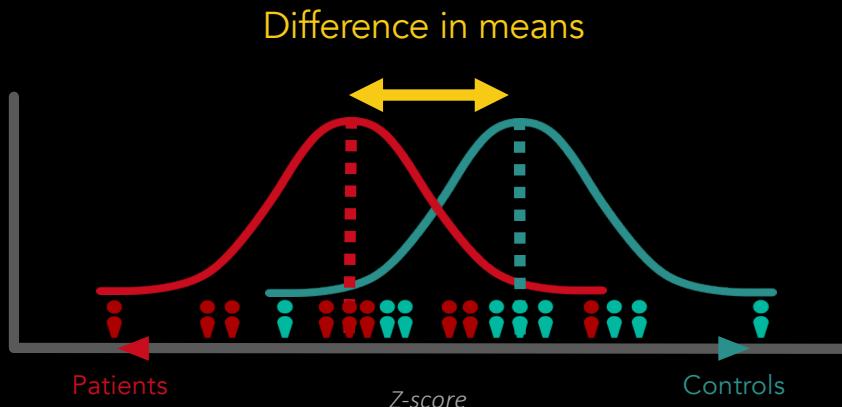
Cortical thickness/ surface area	Covariates (age, sex, parameters site)	Model	Residuals
-------------------------------------	--	-------	-----------

$$\hat{Y}, \sigma_s, \sigma_\epsilon = f(X, \theta) + \epsilon$$

$$Z = \frac{\hat{Y} - Y}{\sqrt{\sigma_s + \sigma_\epsilon}}$$

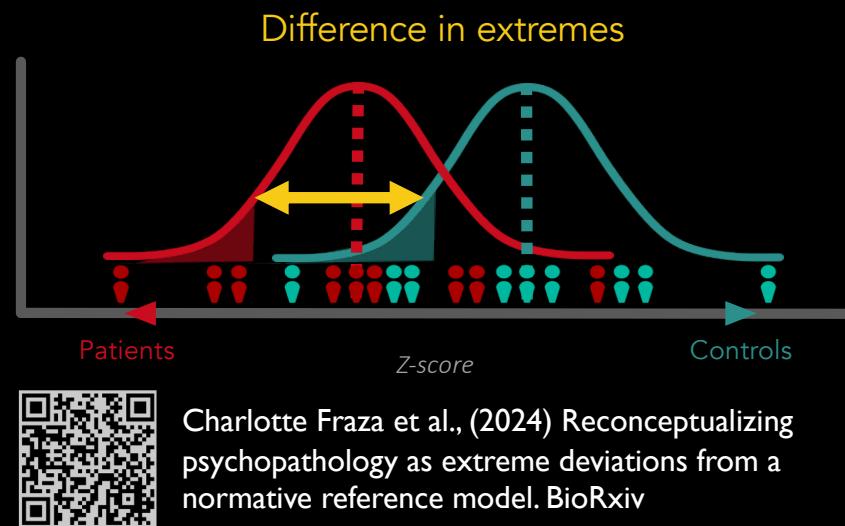
HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS: Structural Imaging



Reminiscent of case-control design, but controlling for individual variation

? Do patients show overlapping deviation scores in brain regions significantly different from control group ?



Charlotte Fraza et al., (2024) Reconceptualizing psychopathology as extreme deviations from a normative reference model. BioRxiv

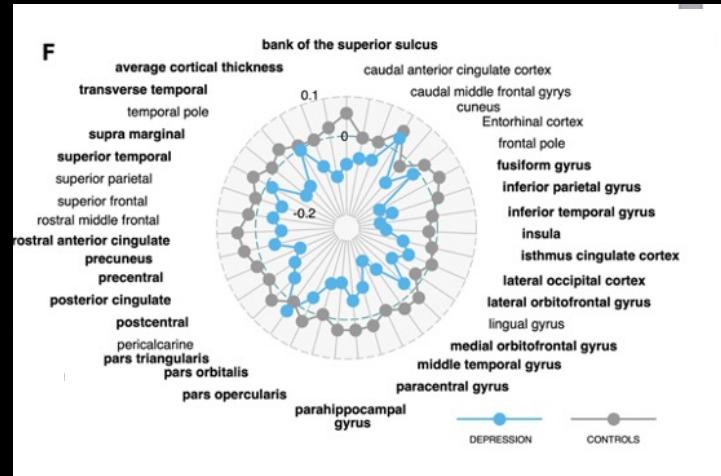
? Are patients more likely to lie in the tails of distribution of (cortical thickness) ?

HETEROGENEITY NORMATIVE MODELLING

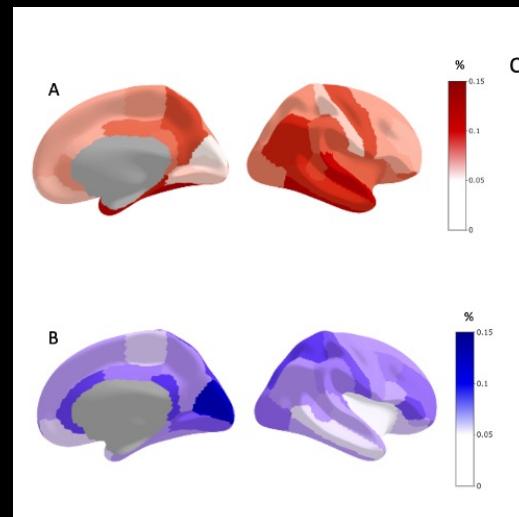
APPLICATIONS: Structural Imaging

I.

Parse heterogeneity



Extreme negative deviations



Extreme positive deviations

2.

Brain-behavior mappings

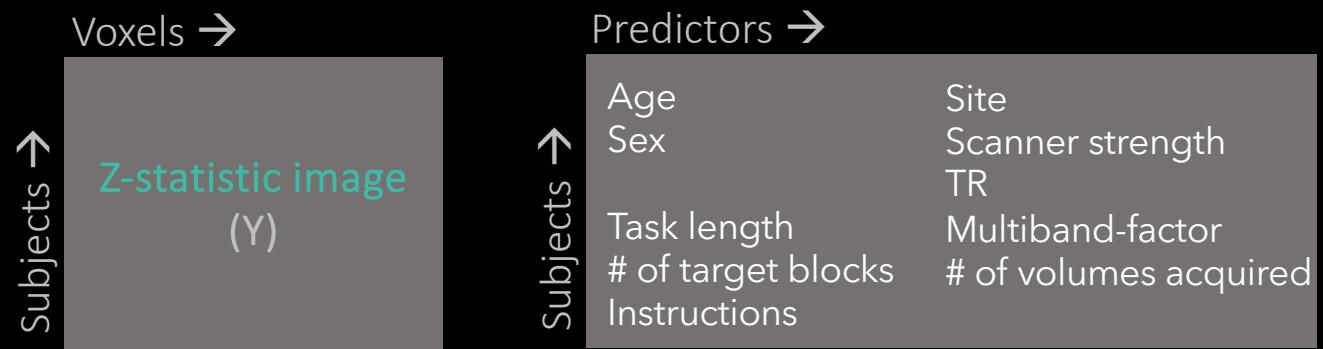
- earlier onset of depression
 - more severe depressive symptoms
 - higher BMI
-
- being remitted
 - not taking antidepressants
 - less severe symptoms

- Training and evaluating the normative model on local dataset

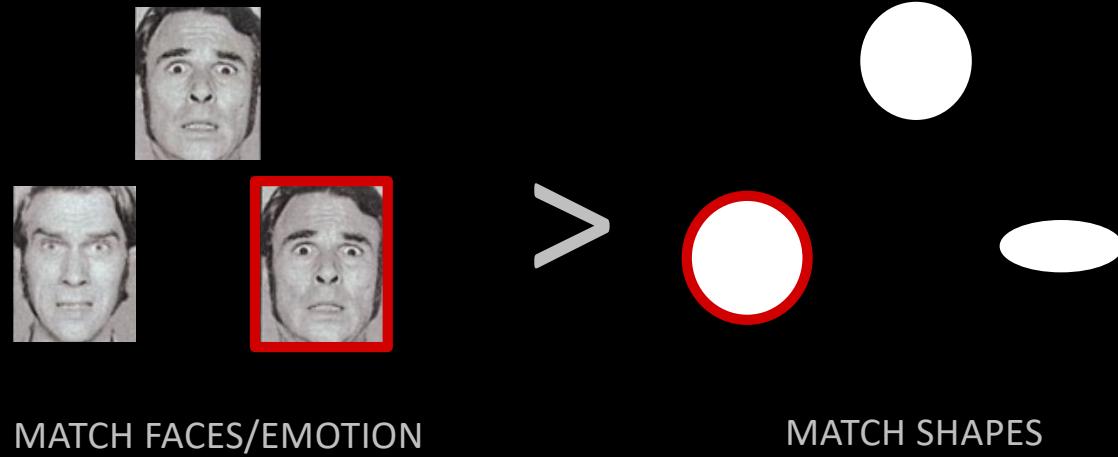
$$Y = f(X, \theta) + \epsilon$$

Legend:

- Bold Signal**: Covariates (age, sex, Task parameters)
- Covariates**: Model parameters
- Model**: Residuals
- Residuals**



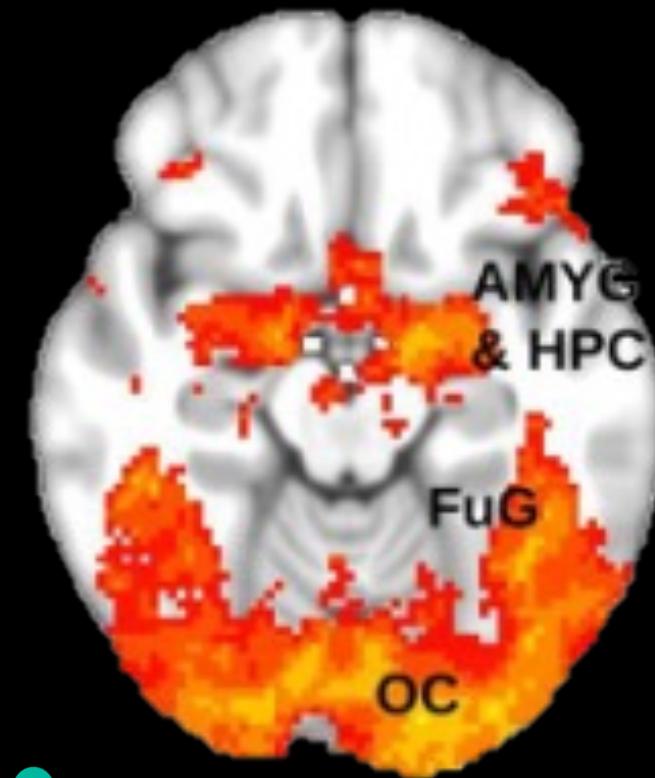
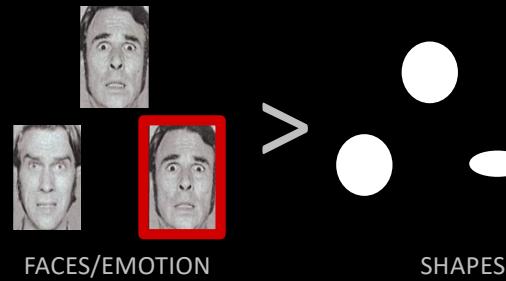
- Training and evaluating the normative model on local dataset

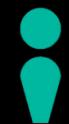


HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset

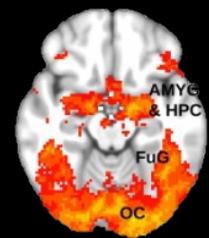


 Unaffected controls

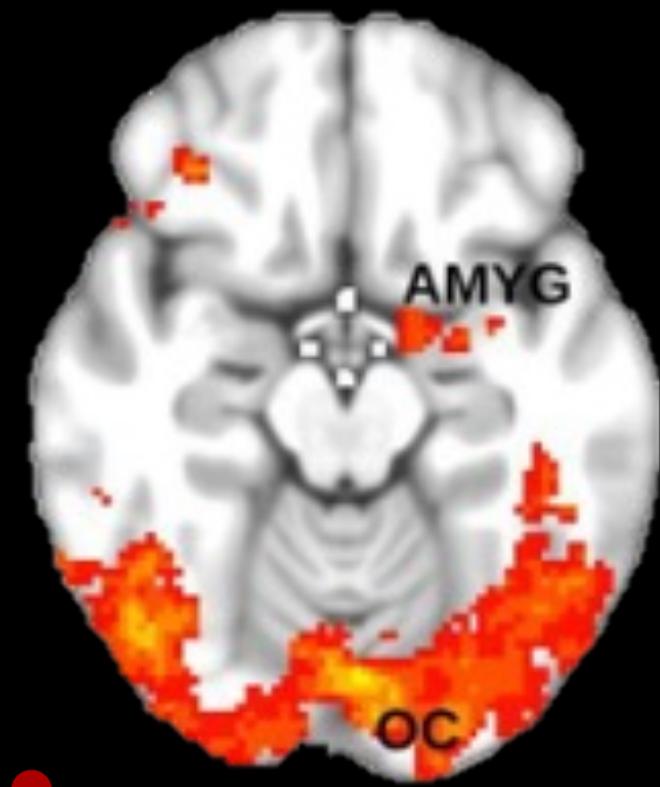
HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset



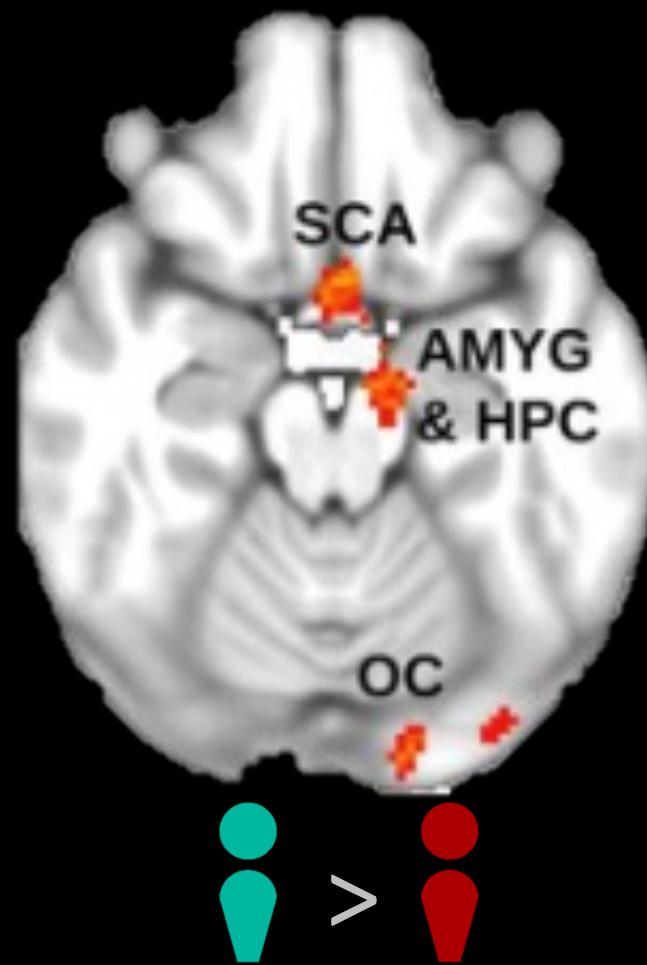
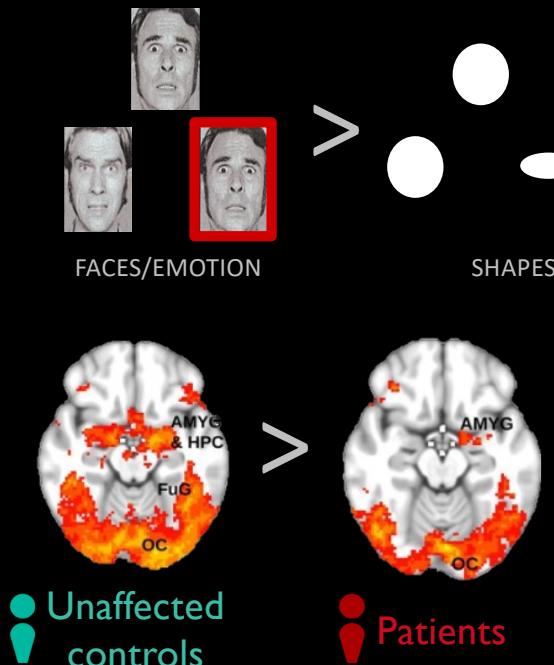
● Unaffected
controls



HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset

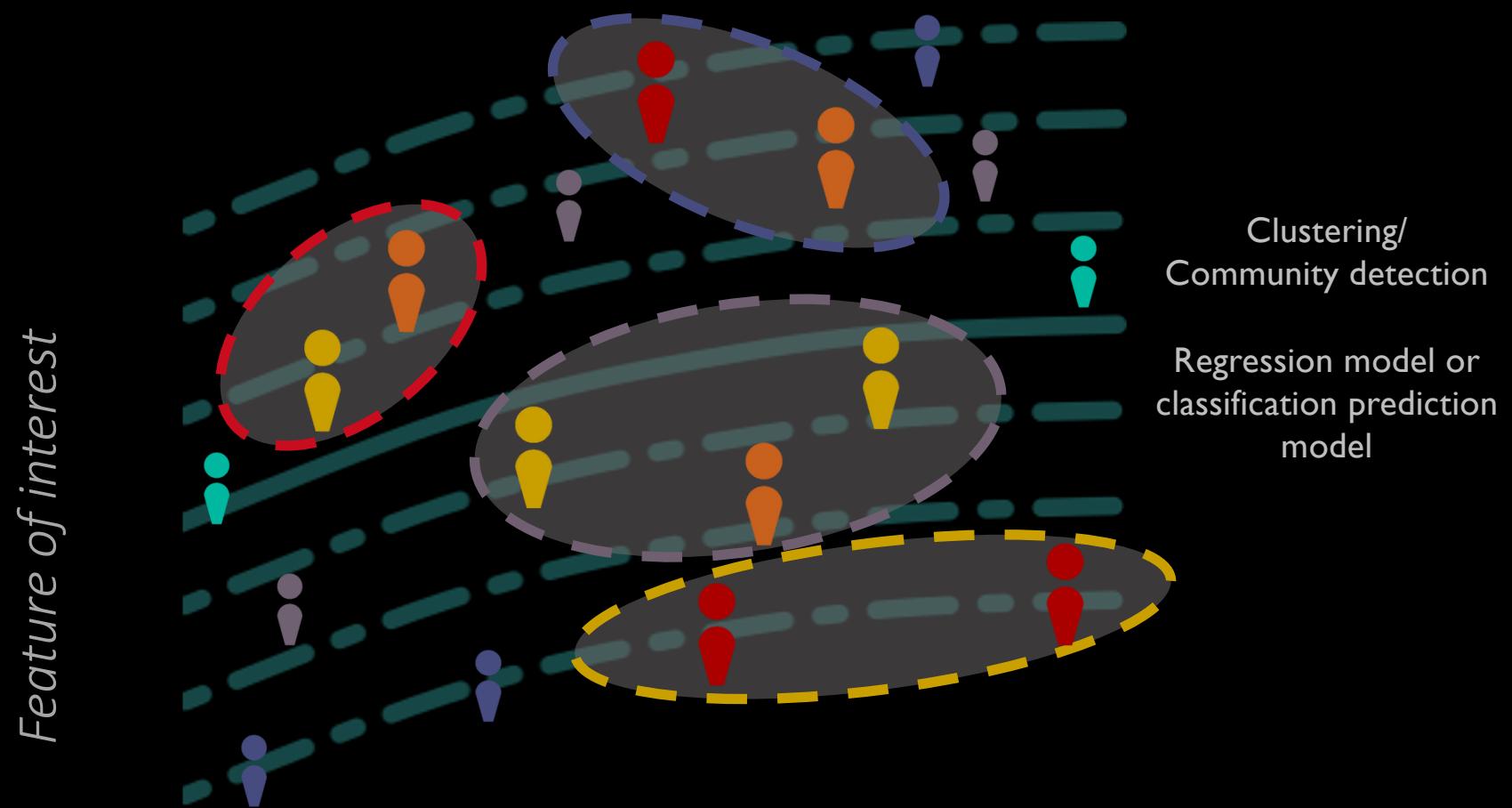


Viering et al (2021) doi: 10.1007/s00787-021-01809-3

HETEROGENEITY NORMATIVE MODELLING

NEUROBIOLOGICAL SUBTYPING

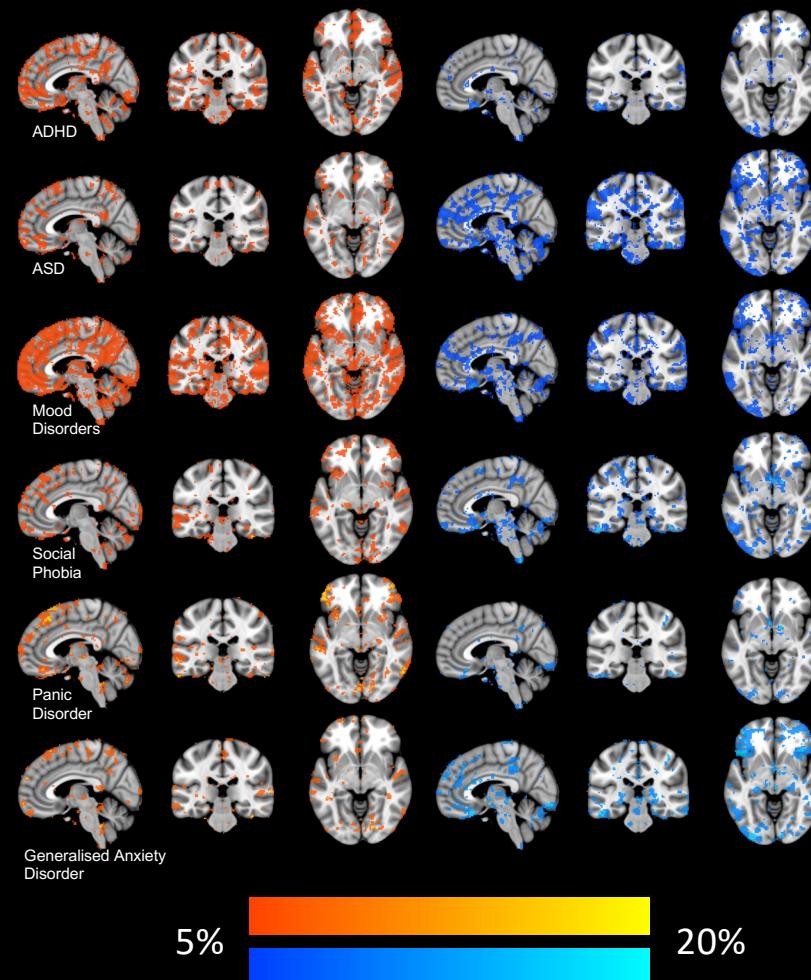
APPLICATIONS



HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset



Hannah Savage et al., (2024) Dissecting task-based fMRI activity using normative modelling: an application to the Emotional Face Matching Task. *Communications Biology* 7.1: 888.

HETEROGENEITY NORMATIVE MODELLING

APPLICATIONS & DEVELOPMENT

Training your
own models

Applying pre-
trained models

Methodology

Training your
own models

Applying pre-
trained models

Methodology

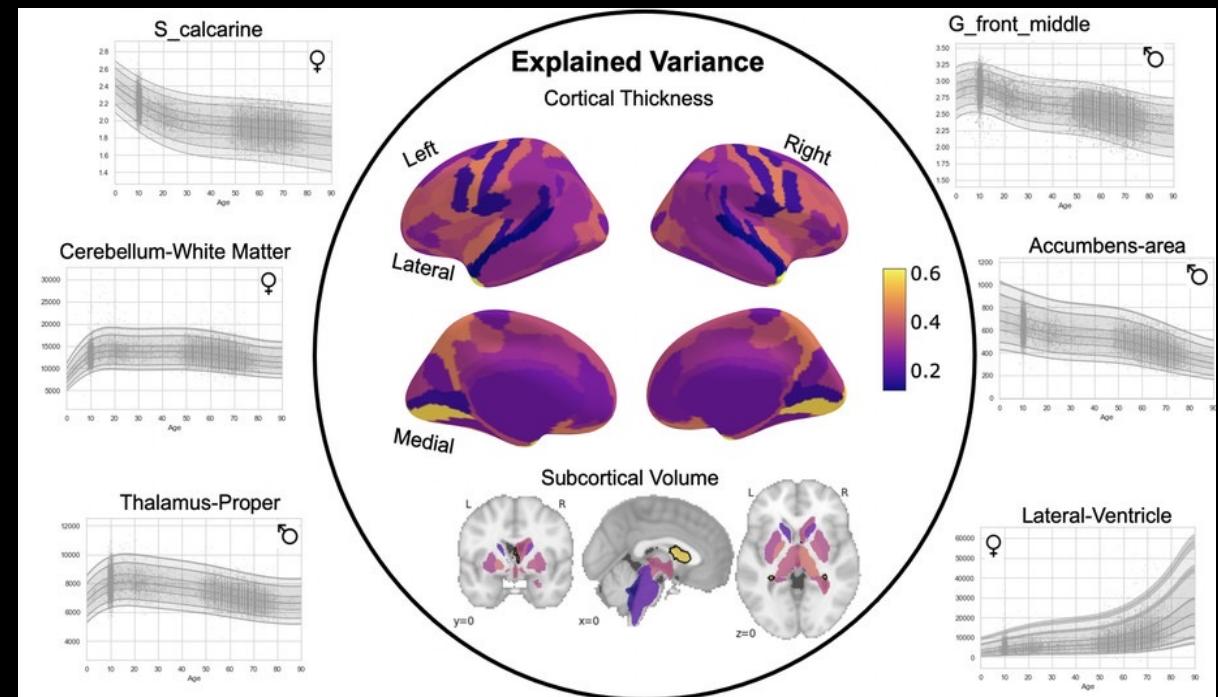
HETEROGENEITY NORMATIVE MODELLING

USE PRETRAINED MODELS

- Using pre-trained models, evaluating on local datasets

- 58,836 individuals
- 82 scan sites
- aged 2–100

Normative models for cortical thickness and subcortical volumes derived from Freesurfer



Saige Rutherford et al.,
(2022) Charting brain growth
and aging at high spatial
precision. *eLife* 11:e72904.

$$\hat{Y}, \sigma_S, \sigma_\epsilon = f(X, \theta) + \epsilon$$

USE PRETRAINED MODELS

- Using pre-trained models, evaluating on local datasets

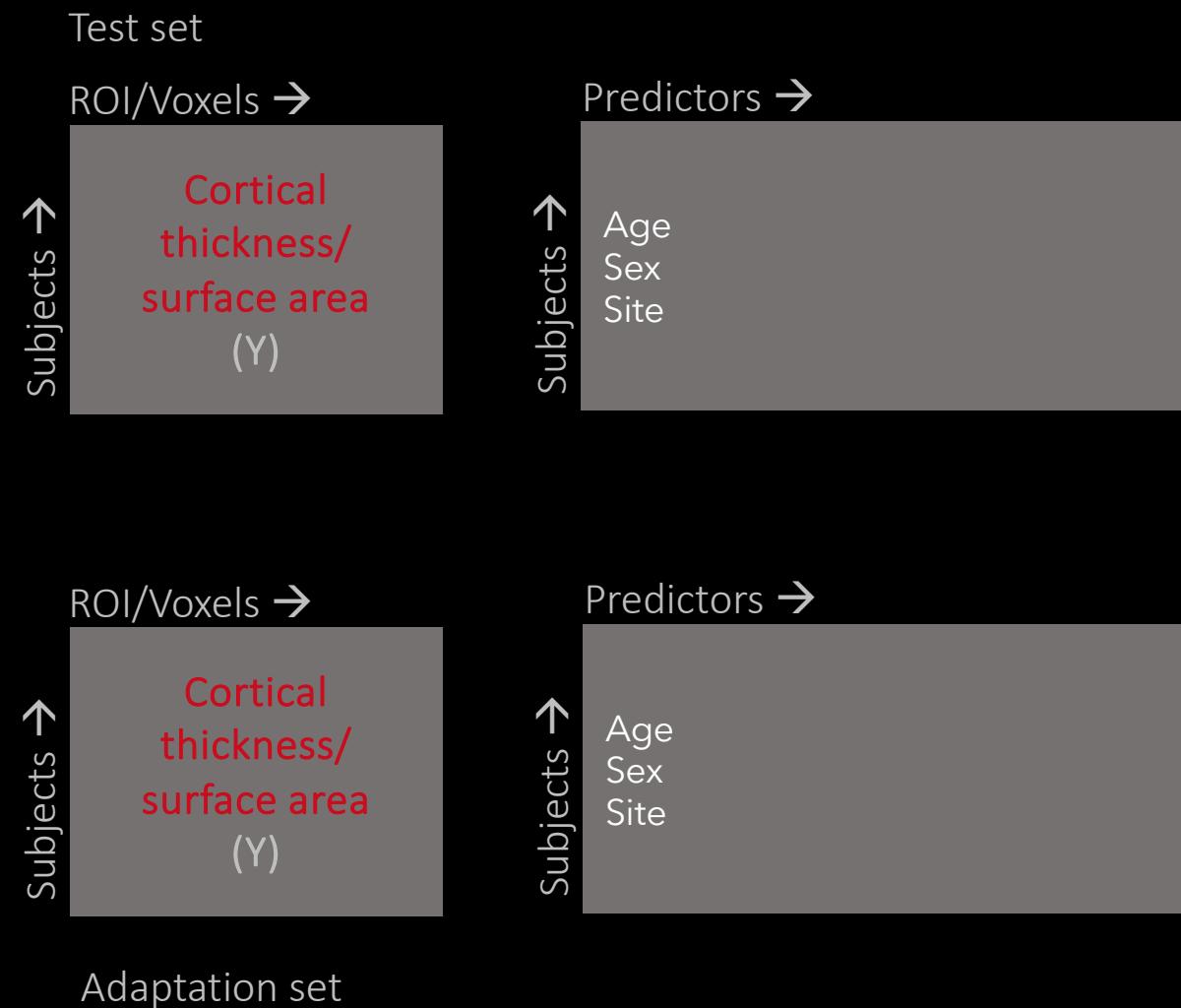
$$Z = \frac{\hat{Y} - Y}{\sqrt{\sigma_s + \sigma_\epsilon}}$$

HETEROGENEITY NORMATIVE MODELLING

USE PRETRAINED MODELS

- Using pre-trained models, evaluating on local datasets

$$\hat{Y} - Y$$



Training your
own models

Applying pre-
trained models

Methodology

Training your
own models

Applying pre-
trained models

Methodology

METHODOLOGICAL ADVANCES

$$Y = f(X, \theta) + \epsilon$$

Brain
(BOLD signal in
voxel, ROI)

Covariates
(age, sex,
task parameters)

Model
parameters

Residuals

Gaussian
process
regression



Warped
Bayesian
linear
regression



Generalized
additive models
of location scale
and shape

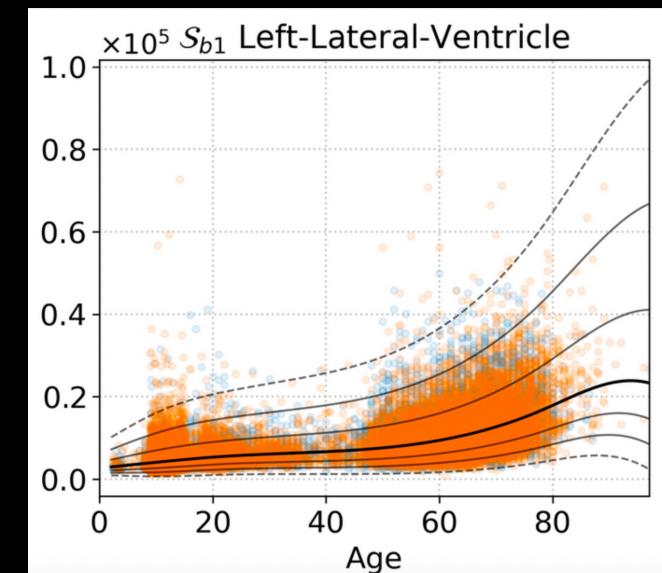
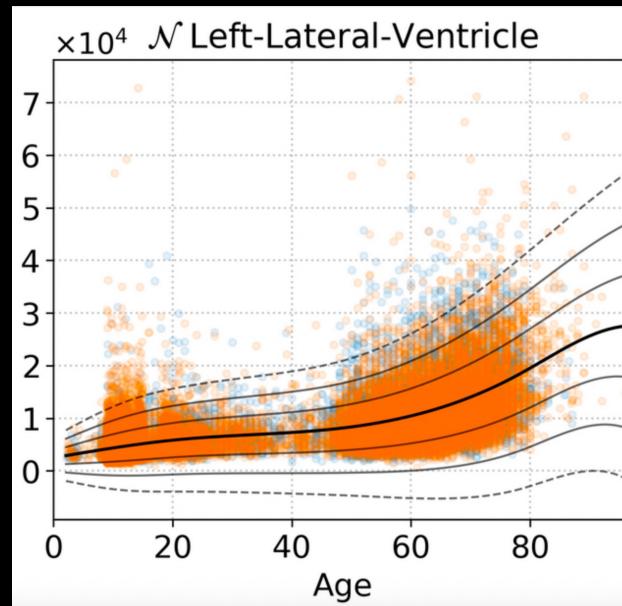


Hierarchical
Bayesian
regression



COMPLEX PHENOTYPES

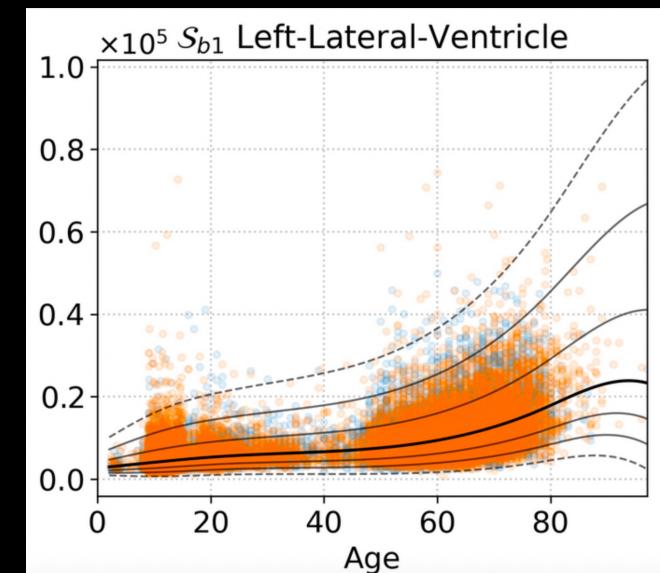
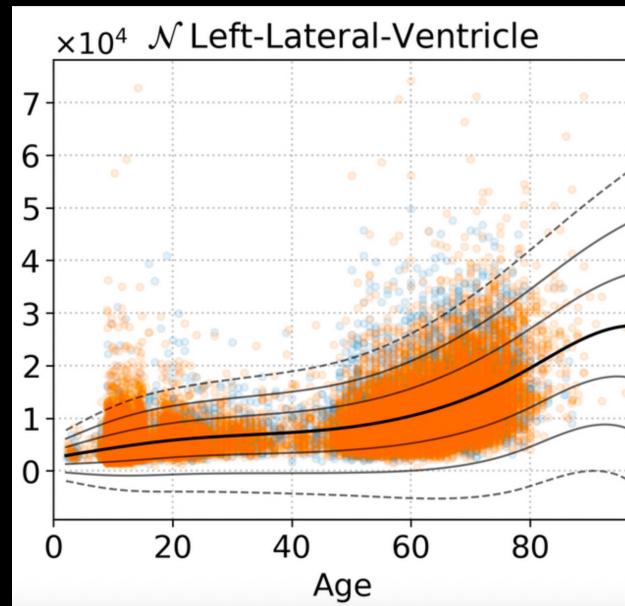
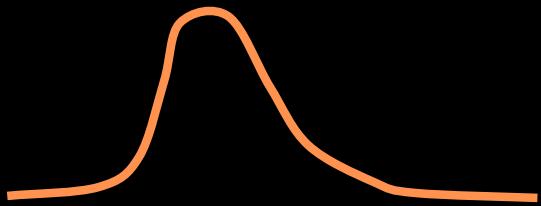
- Hierarchical Bayesian Regression with non-gaussian SHASH likelihood allows to model skewed and heteroscedastic distributions



de Boer, A. A. A. Bayer, J. M. M. et al. (2024) Non-Gaussian normative modelling with hierarchical Bayesian regression. *Imaging Neuroscience* 2, 1–36.

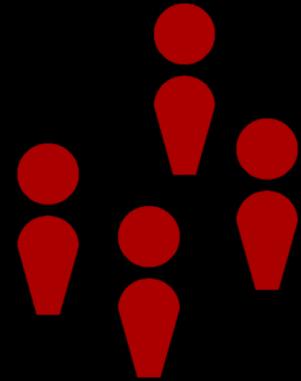
COMPLEX PHENOTYPES

- Hierarchical Bayesian Regression with non-gaussian SHASH likelihood allows to model skewed and heteroscedastic distributions

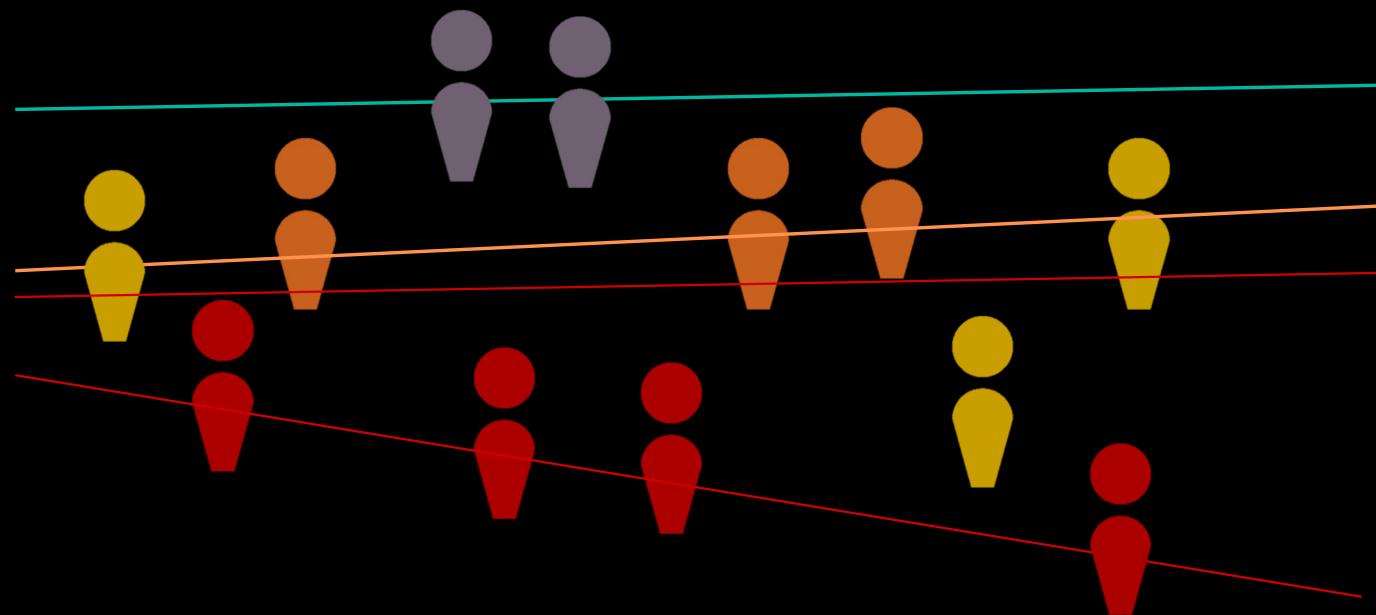


de Boer, A. A. A. Bayer, J. M. M. et al. (2024) Non-Gaussian normative modelling with hierarchical Bayesian regression. *Imaging Neuroscience* 2, 1–36.

SITE EFFECTS



SITE EFFECTS

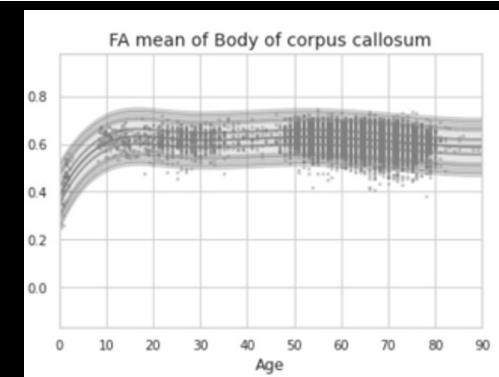
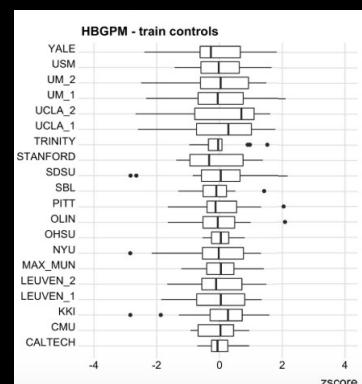
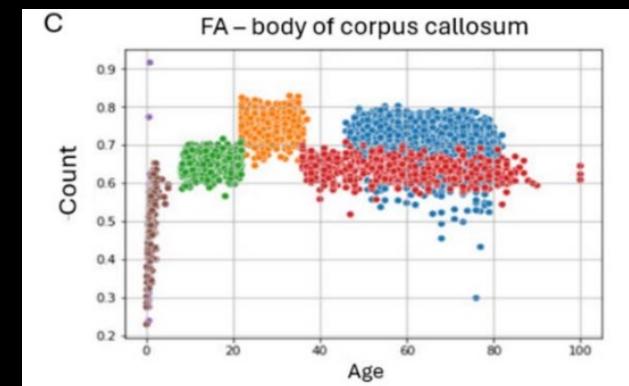
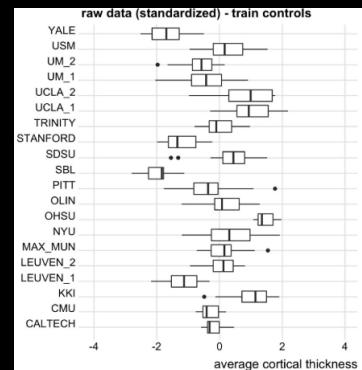


SITE EFFECTS



SITE EFFECTS

- Including site as batch effect into normative model
- Inverting z-scores between sites

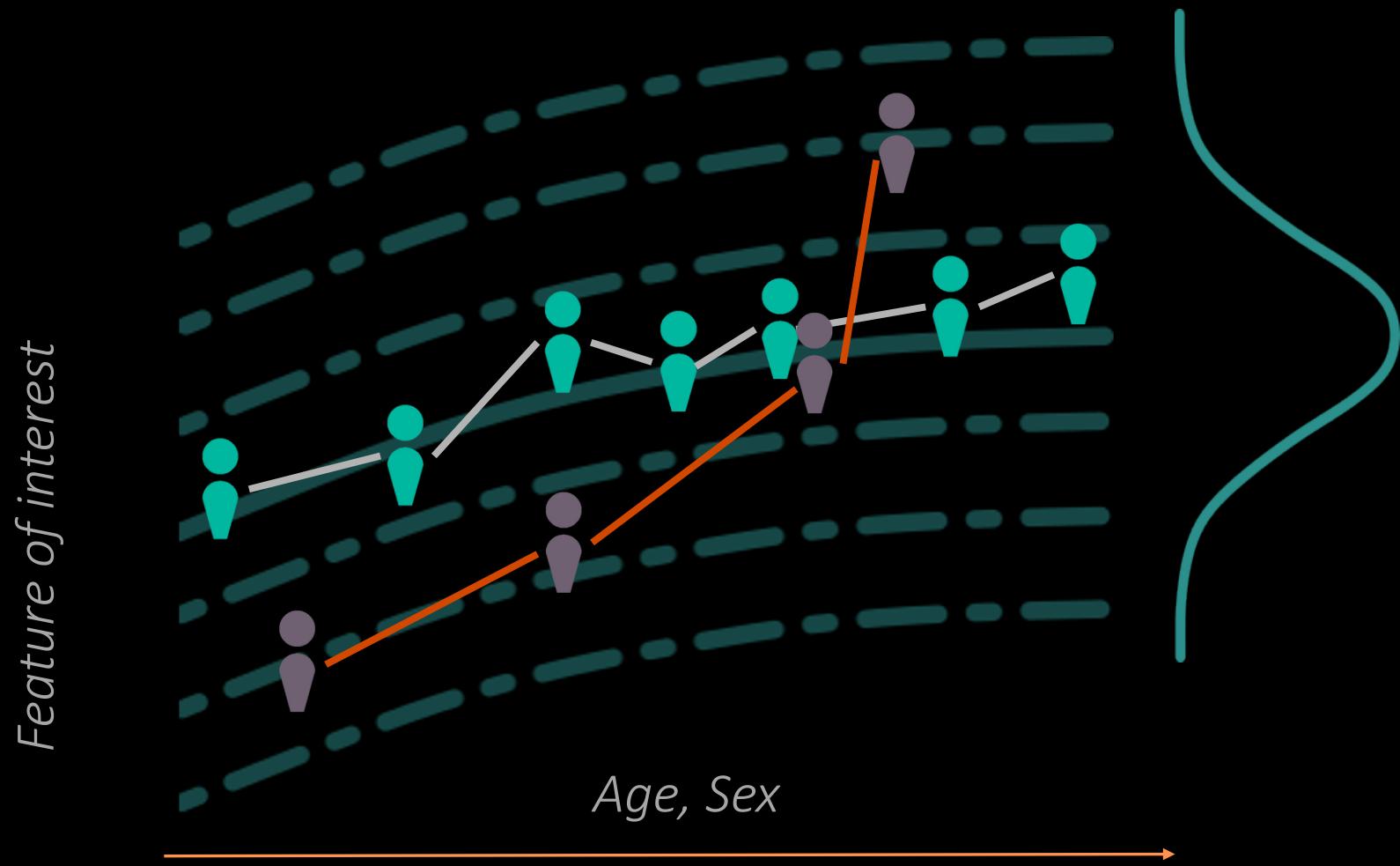


Bayer, J. M. M. *et al.* Accommodating site variation in neuroimaging data using normative and hierarchical Bayesian models. *Neuroimage* **264**, 119699 (2022).

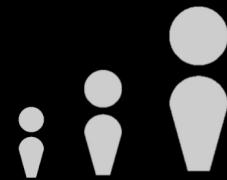


Cirstian, R. *et al.* Lifespan normative models of white matter fractional anisotropy: Applications to Early Psychosis. *bioRxiv.org* 2024.12.11.627897 (2024) doi:10.1101/2024.12.11.627897.

LONGITUDINAL NORMATIVE MODELLING

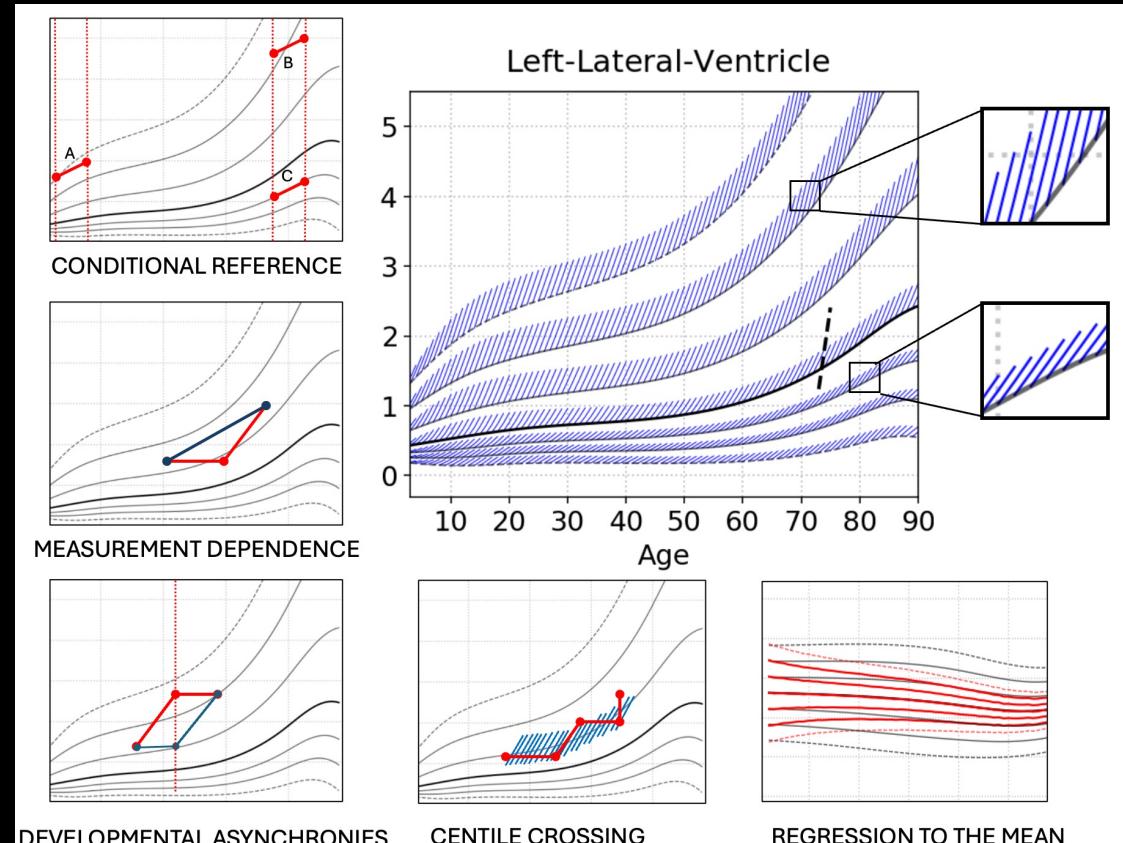


LONGITUDINAL NORMATIVE MODELLING



**Longitudinal models –
From longitudinal data
Thrive lines**

- 22000 longitudinal data points
- Age range: 7 days – 97 years
- Thrive lines: estimate the point of failure to thrive
- Allow for true longitudinal predictions and an estimate of centile crossing
- Normal rate of change varies with age



Bayer 2024, in prep.

https://github.com/likeajumprope/Tutorial_normative_modelling

Tasks

Task 1: Fitting normative models from scratch  [Open in Colab](#)

Task 2: Applying pre-trained normative models  [Open in Colab](#)

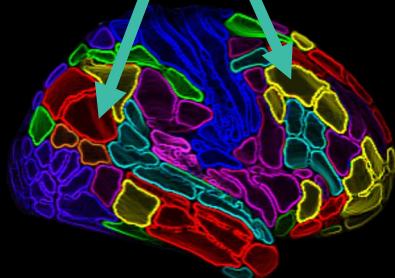
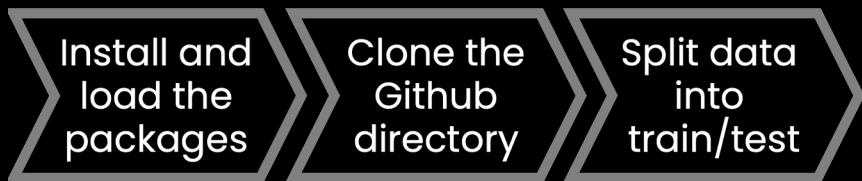
TUTORIAL I.

ESTIMATING LIFESPAN NORMATIVE MODELS

- 
- Install and load the packages
 - Clone the Github directory
 - Split data into train/test

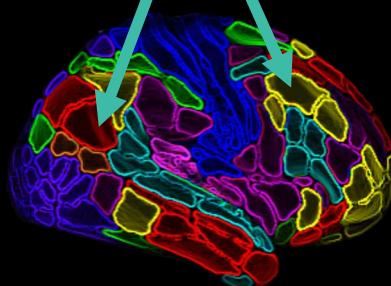
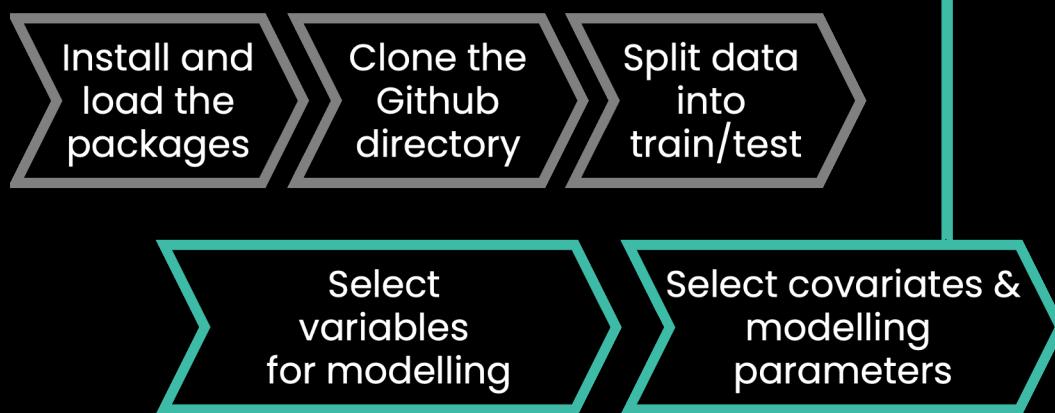
TUTORIAL I.

ESTIMATING LIFESPAN NORMATIVE MODELS

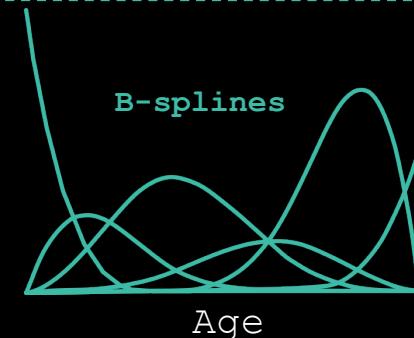


TUTORIAL I.

ESTIMATING LIFESPAN NORMATIVE MODELS



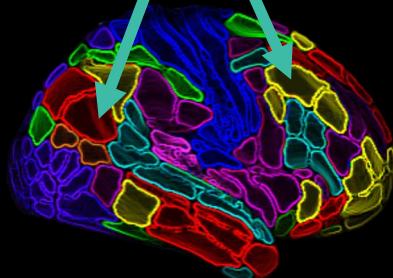
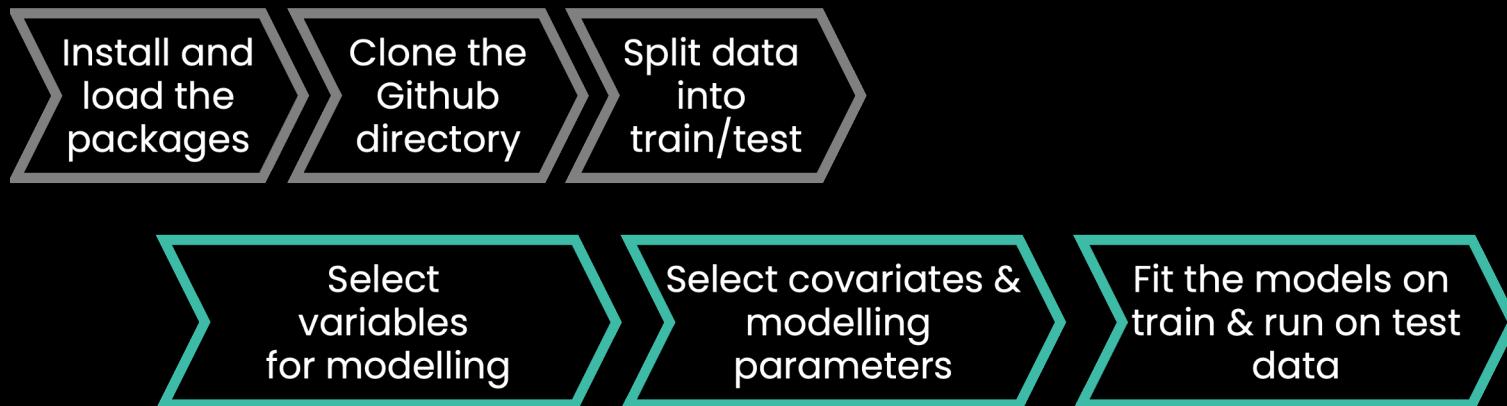
- `["age", "sex"]`
- Warp modelled variable?
- Set-up B-spline
- Set-up outlier threshold



Subjects	Intercept (vector of ones)	Age	Sex
	Dummy coding of site 1		
	Dummy coding of site 2		
	Dummy coding of site 3		
	B-splines ('7 columns)		

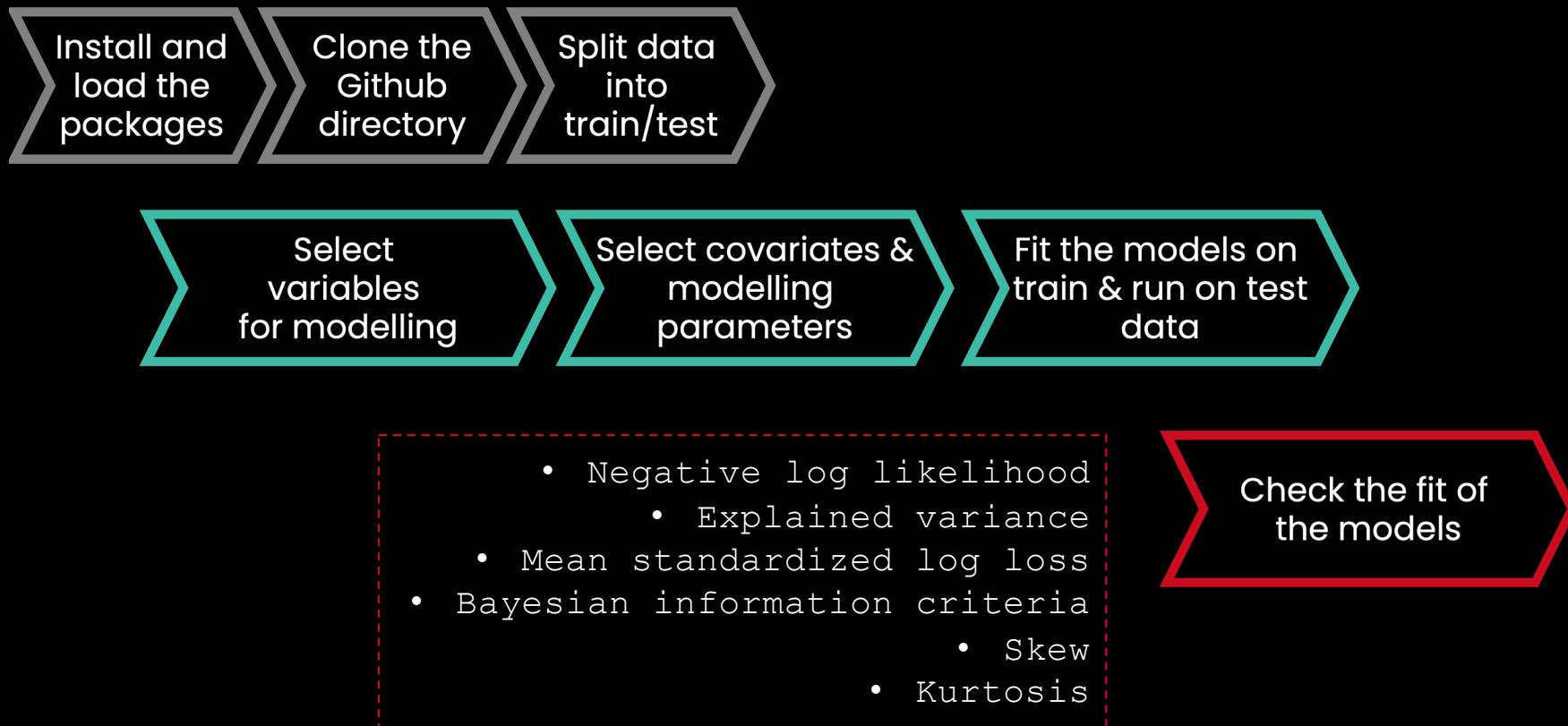
TUTORIAL I.

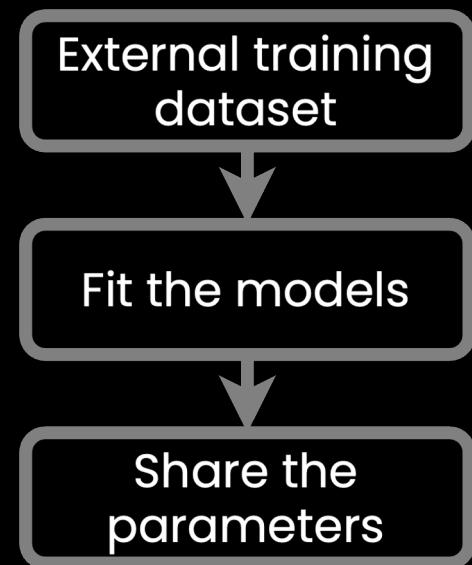
ESTIMATING LIFESPAN NORMATIVE MODELS

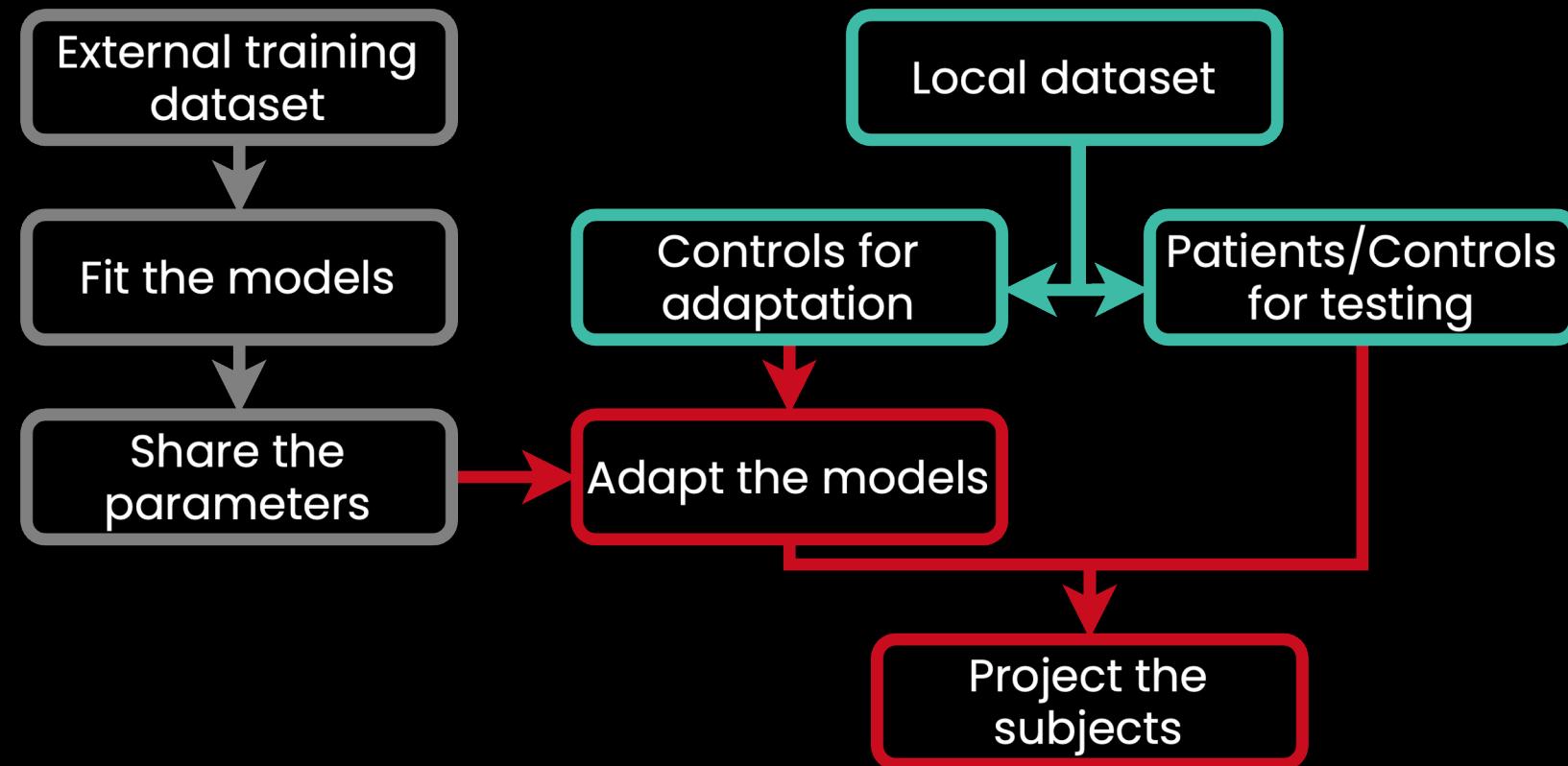


TUTORIAL I.

ESTIMATING LIFESPAN NORMATIVE MODELS







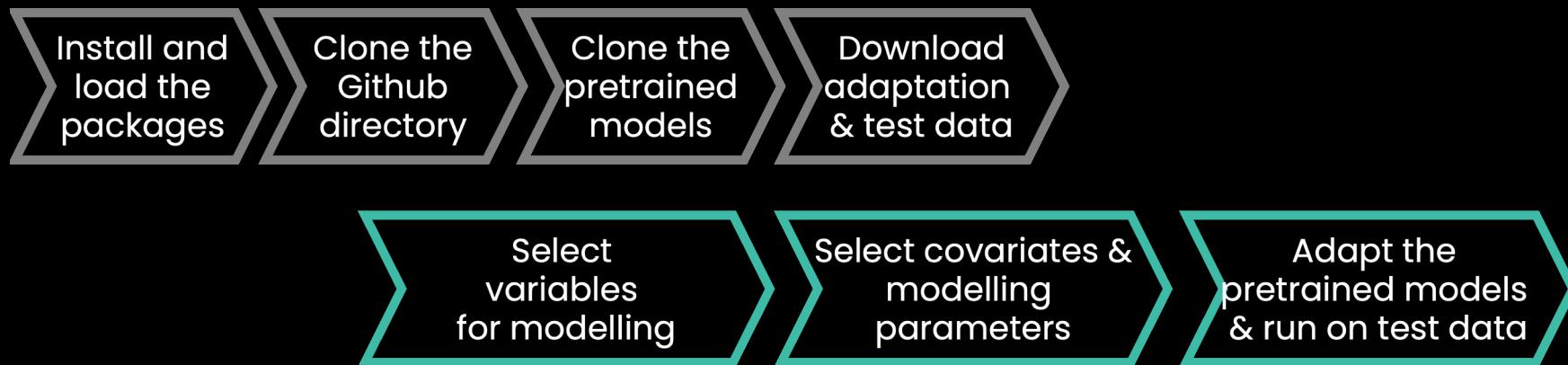
TUTORIAL II.

APPLYING PRE-TRAINED NORMATIVE MODELS



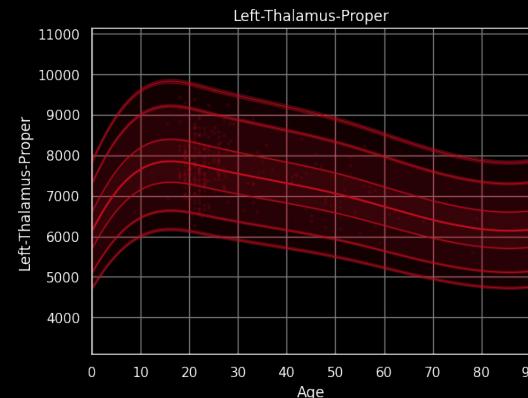
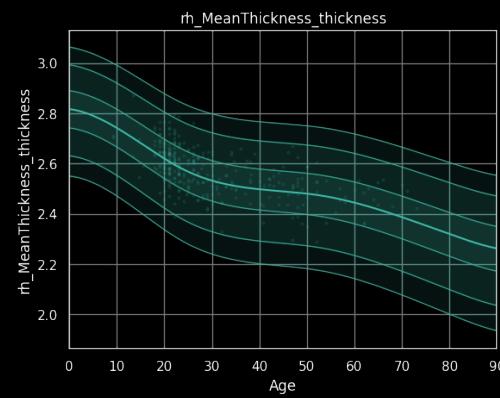
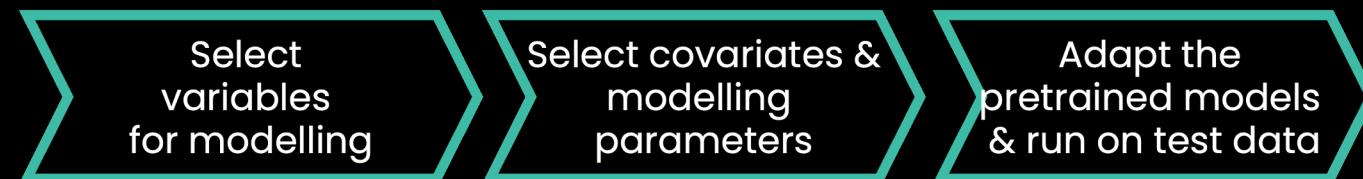
TUTORIAL II.

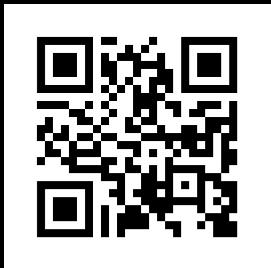
APPLYING PRE-TRAINED NORMATIVE MODELS



TUTORIAL II.

APPLYING PRE-TRAINED NORMATIVE MODELS





Download the toolbox
here:
github.com/amarquand



pcnportal.dccn.nl



<https://pcntoolkit.readthedocs.io>

Predictive Clinical Neuroscience Lab

Professor Andre Marquand



THANK YOU!

https://github.com/likeajumprope/Tutorial_normative_modelling

Johanna.Bayer@radboudumc.nl
Barbora.Rehak-Buckova@radboudumc.nl
Charlotte.Fraza@radboudumc.nl
Hannah.savage@ucl.ac.uk