Normative modelling: the what, when and why

Dr. Johanna Bayer

Post Doctoral Researchers

Predictive Clinical Neuroscience Lab

Predictive Clinical Neuroscience Lab

<u>Slides by:</u> @DrHannahSavage Dr. Charlotte Fraza @CharFraza Dr. Barbora Rehák Bučková @BarboraRehak

OUTLINE

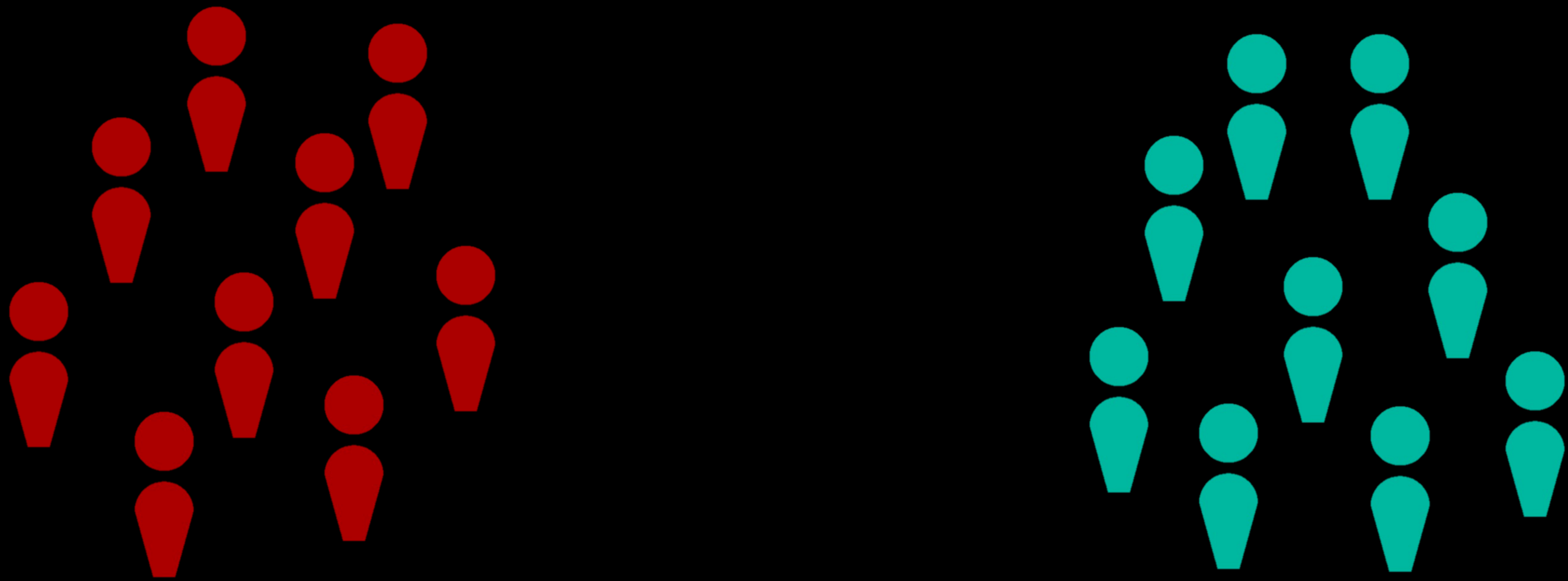
1. Short and conceptual introduction to normative modelling
2. How to train your own normative models
3. How to use out pretrained models
4. Methodological developments
5. Tutorial (Jupyter notebooks)

Resources:

- PCN Toolkit: https://pcntoolkit.readthedocs.io/en/latest/
- PCN Portal: https://pcnportal.dccn.nl/
- Tutorials and Repository for this talk: https://github.com/likeajumprope/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/PCNtoolkit-demo

THE SHORTCOMINGS OF CLASSICAL STATISTICS

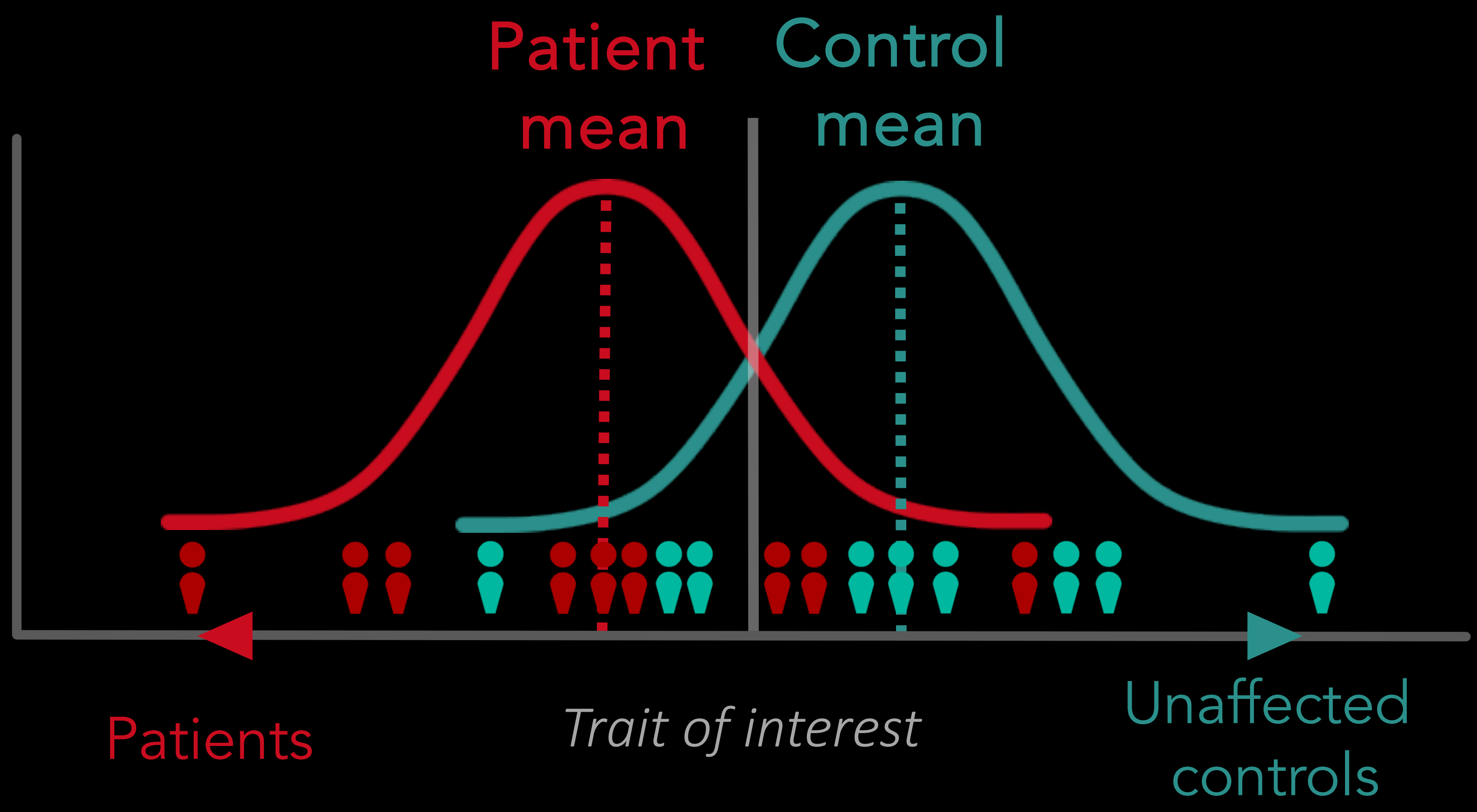
CASE CONTROL



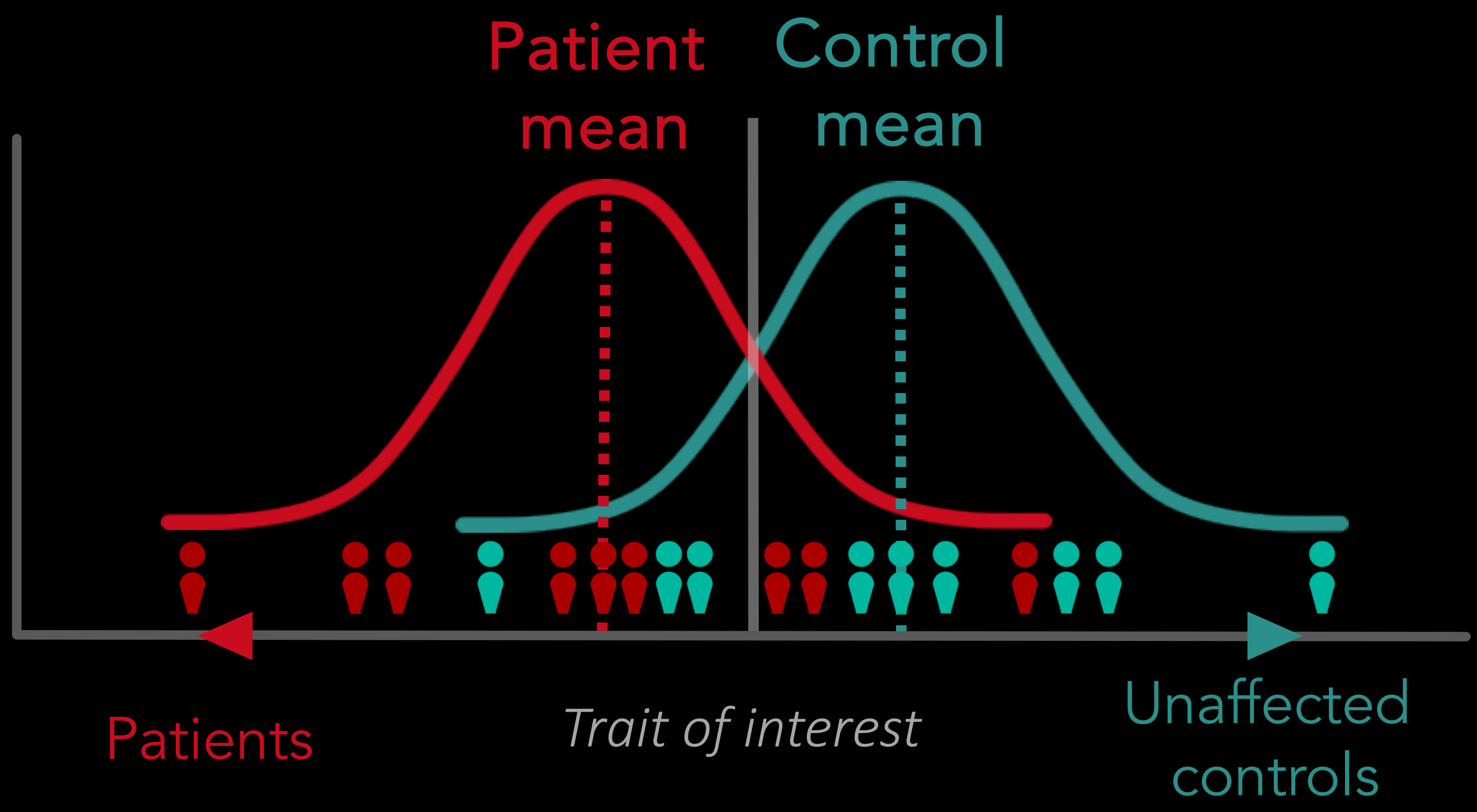
Patients

Unaffected controls

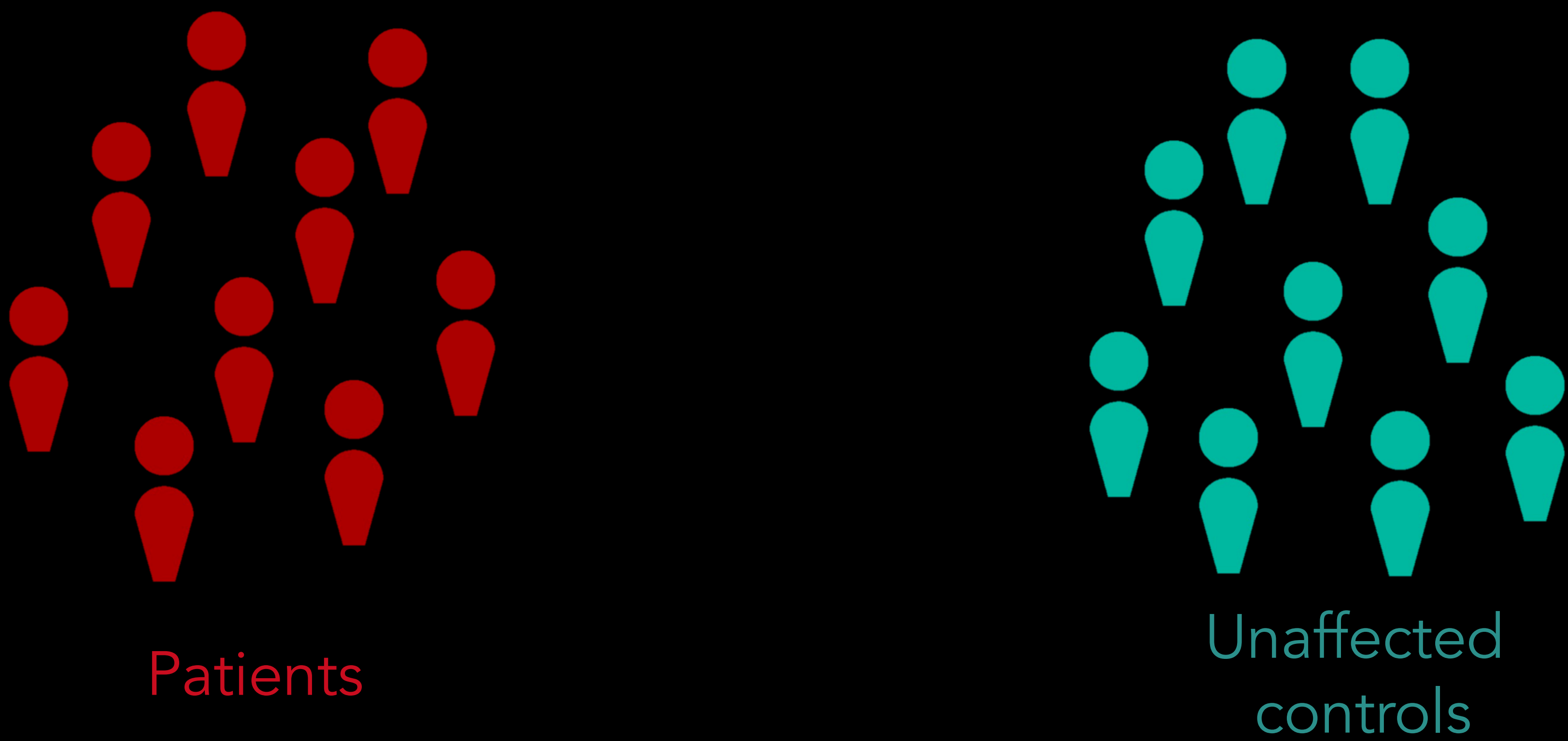
THE SHORTCOMINGS OF CLASSICAL STATISTICS



THE SHORTCOMINGS OF CLASSICAL STATISTICS



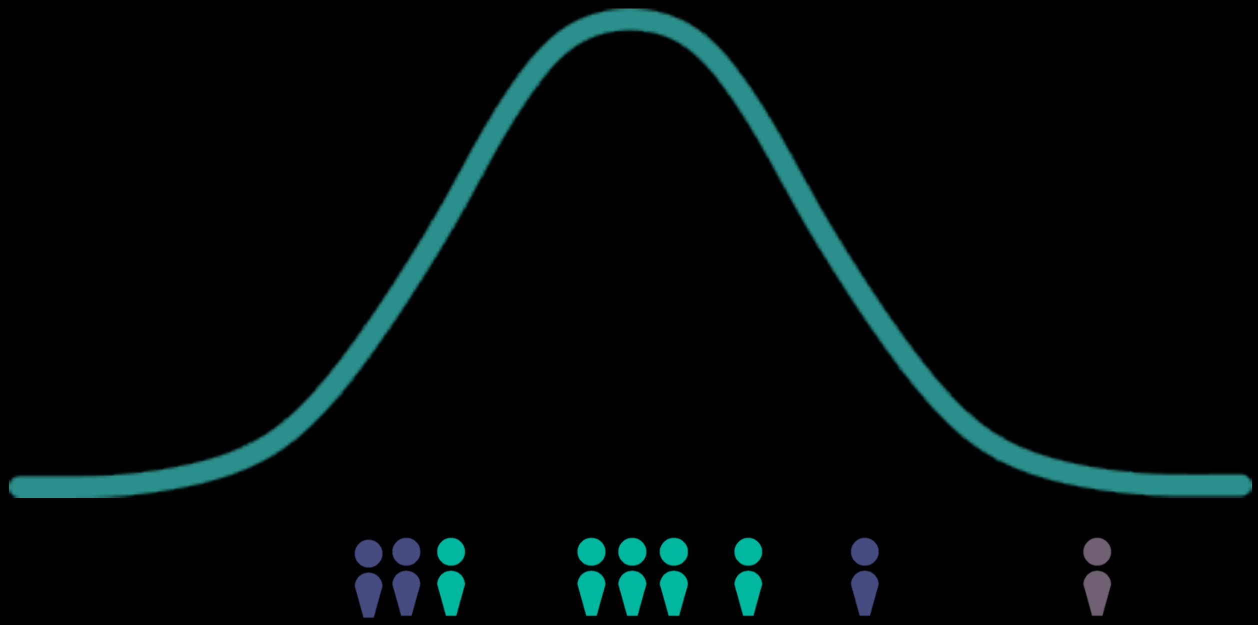
ON HETEROGENEITY



ON HETEROGENEITY

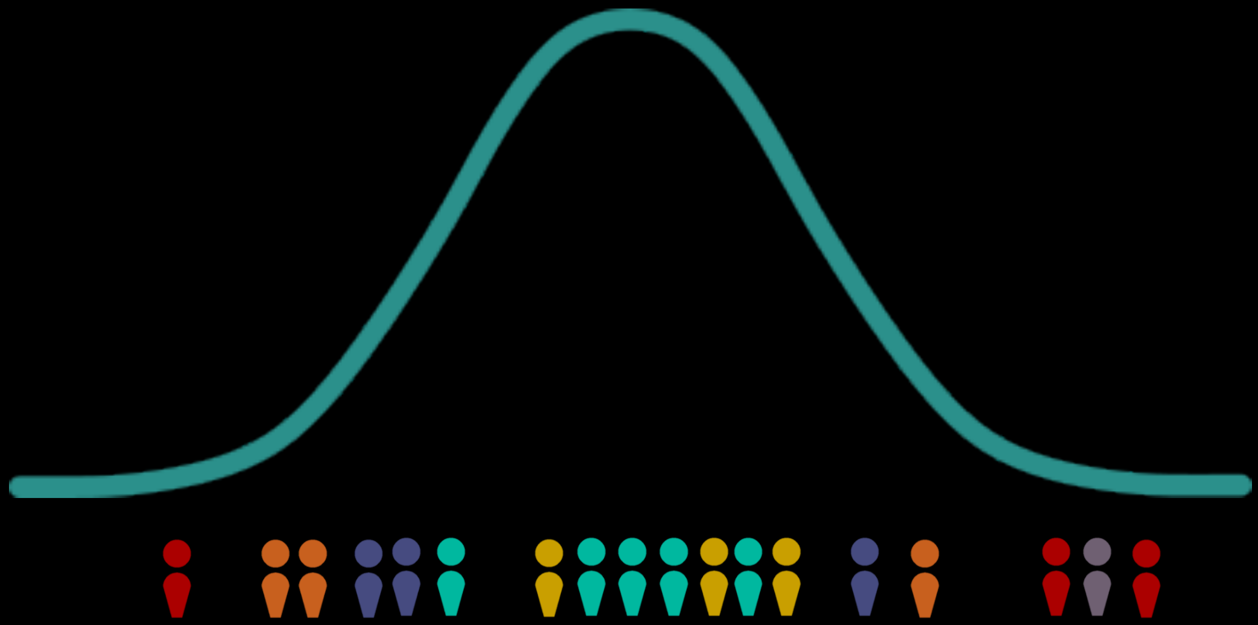
ON HETEROGENEITY

NORMATIVE MODELLING



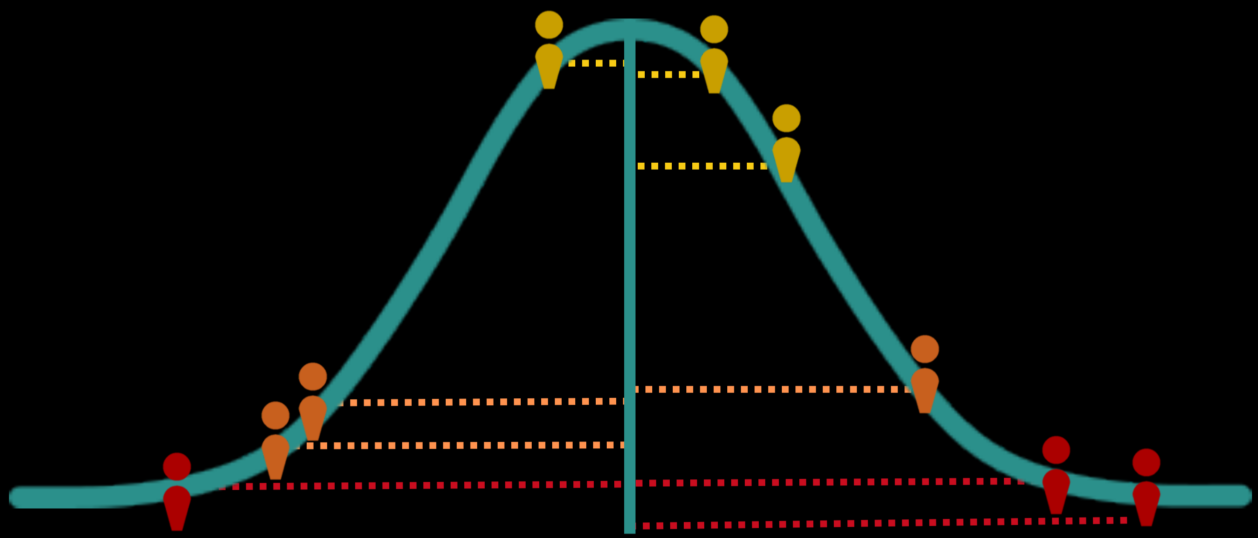
HETEROGENEITY

NORMATIVE MODELLING



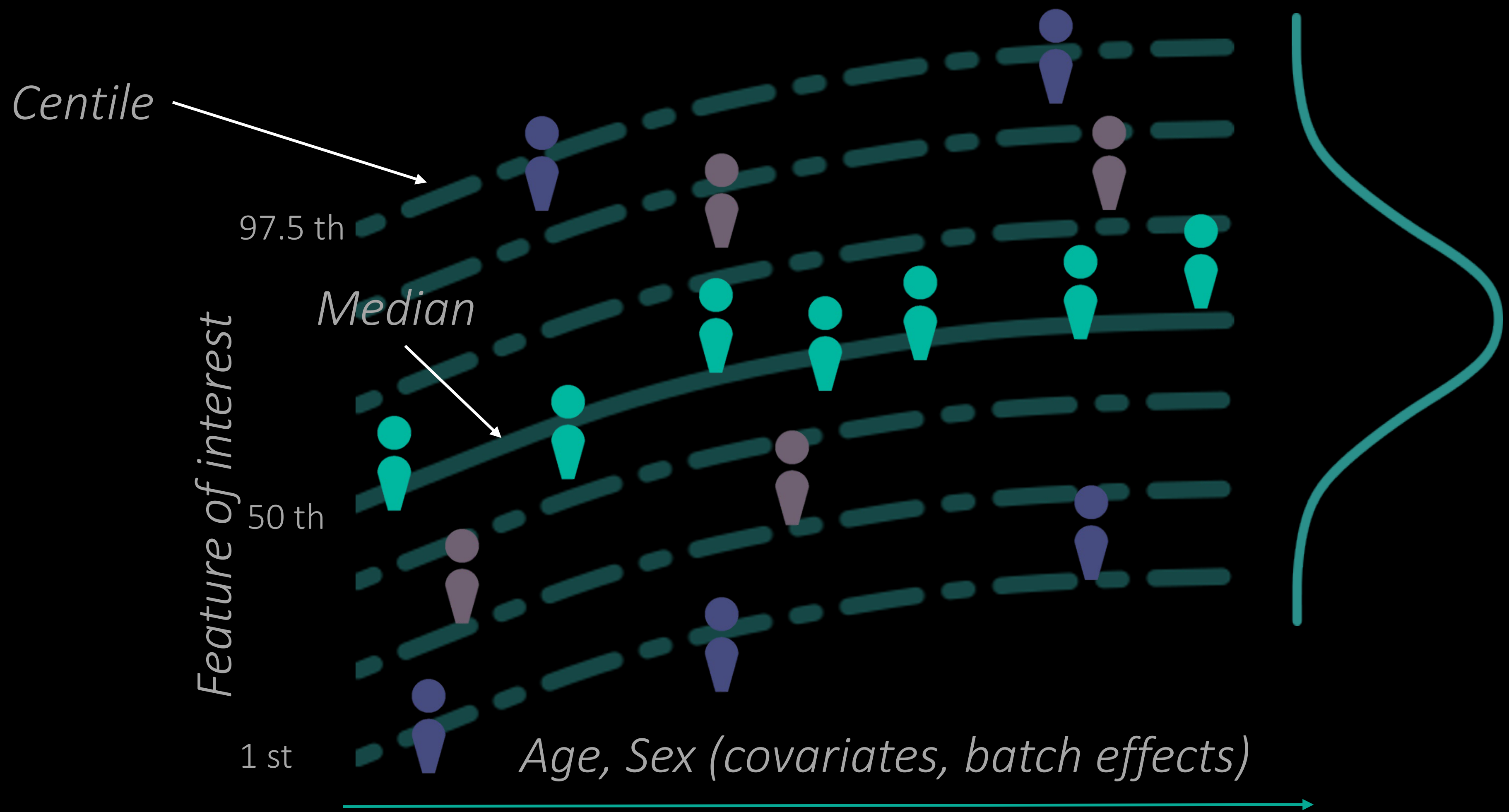
HETEROGENEITY

NORMATIVE MODELLING

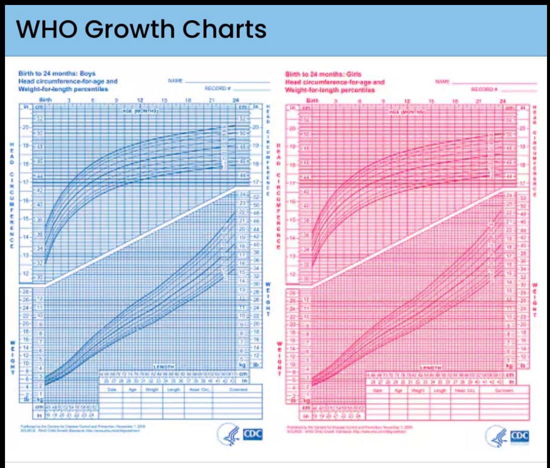


HETEROGENEITY

NORMATIVE MODELLING

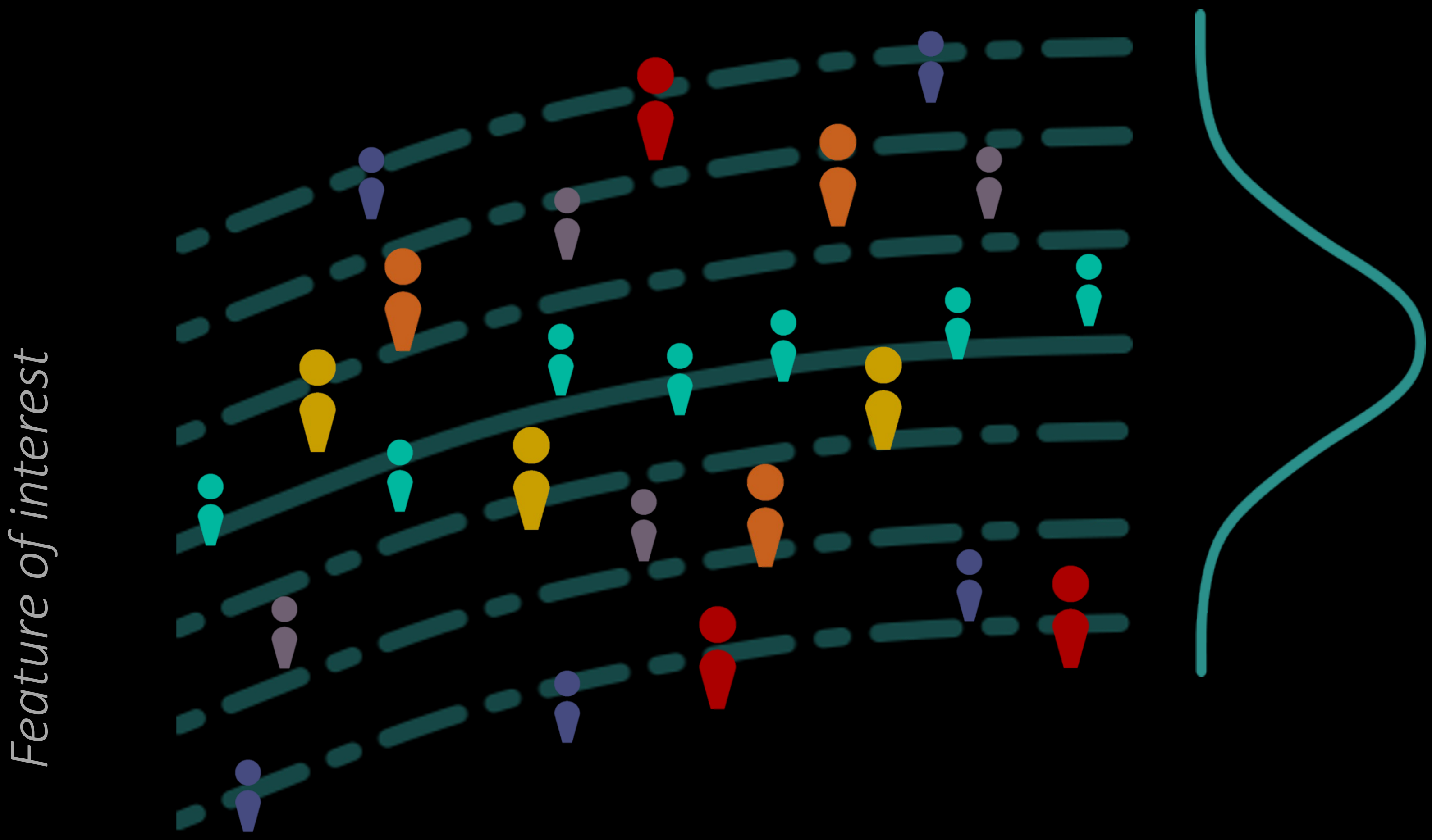


NORMATIVE MODELLING

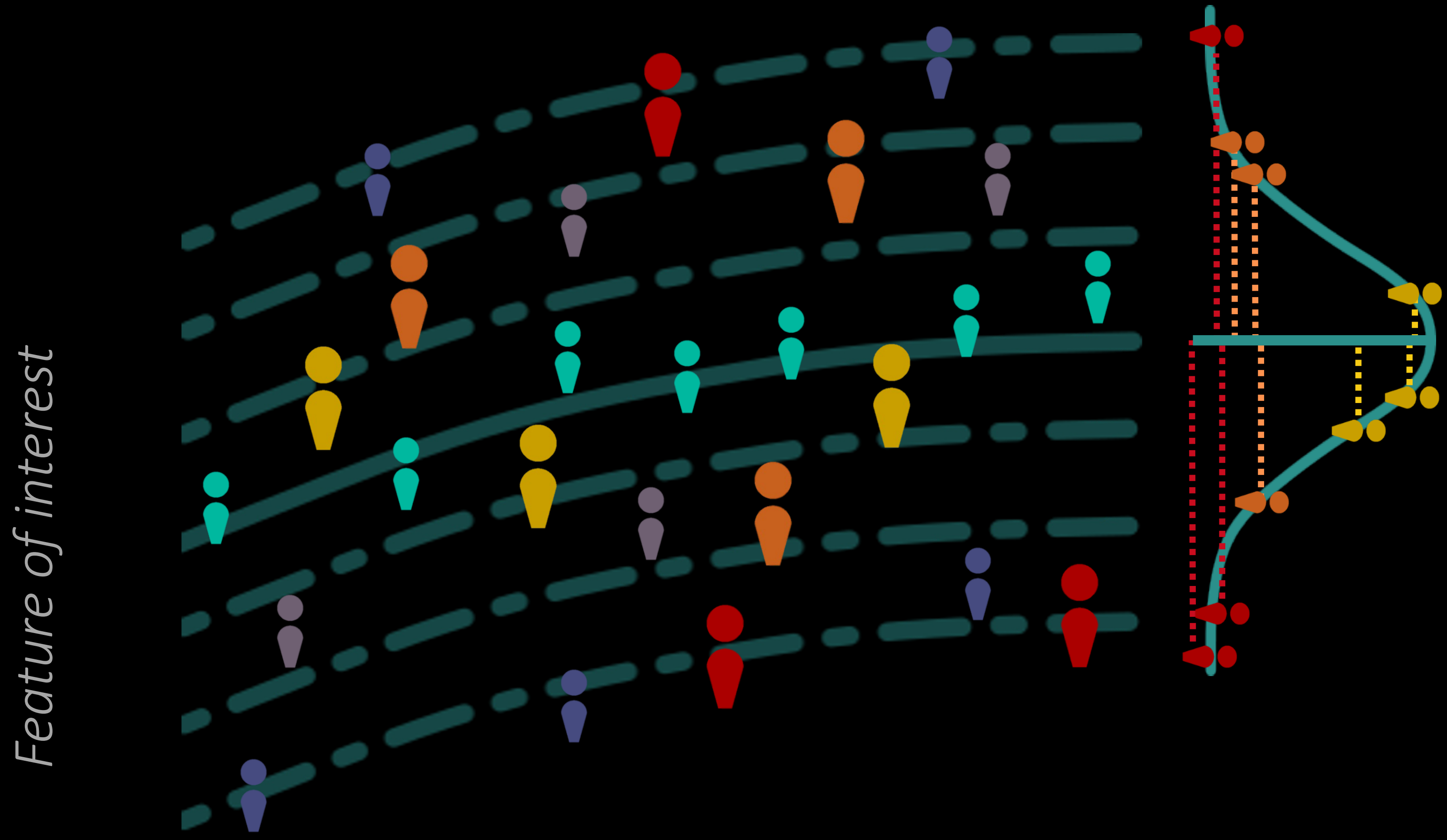


HETEROGENEITY

NORMATIVE MODELLING

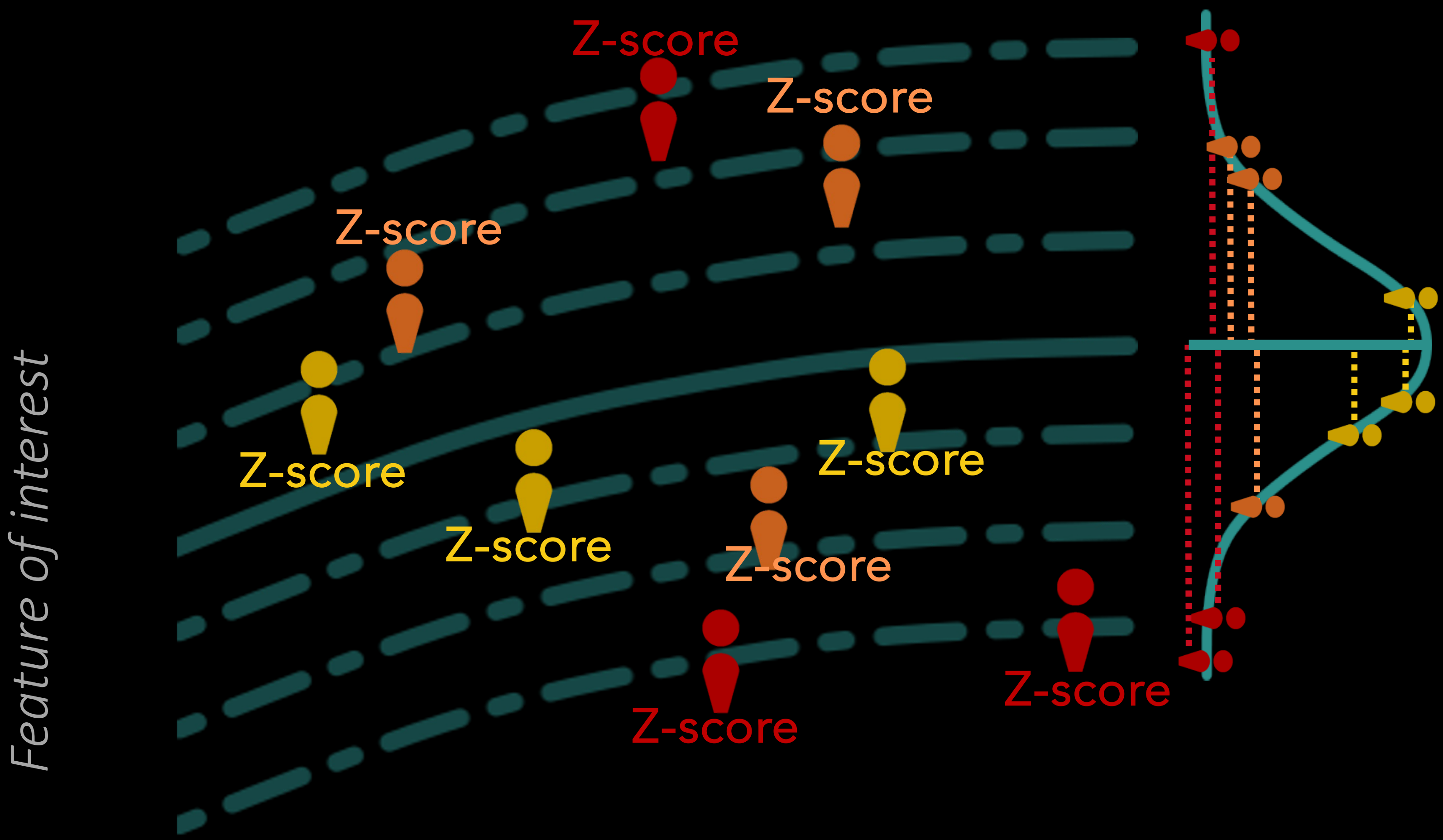


NORMATIVE MODELLING



HETEROGENEITY

NORMATIVE MODELLING



APPLICATIONS & DEVELOPMENT

Training your own models

Applying pre-trained models

Methodology

APPLICATIONS & DEVELOPMENT

Training your own models

Applying pre-trained models

Methodology

APPLICATIONS & DEVELOPMENT

Y

Brain
(BOLD signal in voxel, ROI)

APPLICATIONS & DEVELOPMENT

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

Brain
(BOLD signal in voxel, ROI)

Covariates
(age, sex, task parameters)

Model parameters

Residuals

APPLICATIONS & DEVELOPMENT

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

Brain (BOLD signal in voxel, ROI)

Covariates (age, sex, task parameters)

Model parameters

Residuals

Imaging measures →

Subjects ↑

Responses (Y)

Predictors →

Subjects ↑

Covariates (X)

- Training and evaluating the normative model on local dataset

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

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

- Training and evaluating the normative model on local dataset

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

Cortical thickness/ surface area

Covariates (age, sex, site)

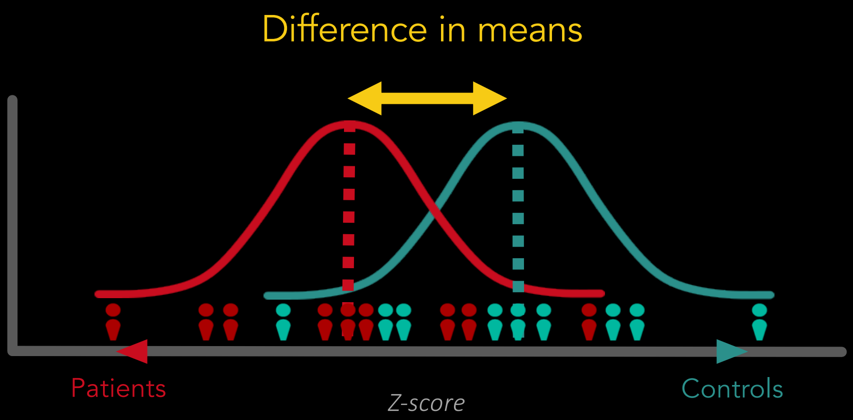
Model parameters

Residuals

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

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

APPLICATIONS: Structural Imaging



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

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

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

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

APPLICATIONS: Structural Imaging

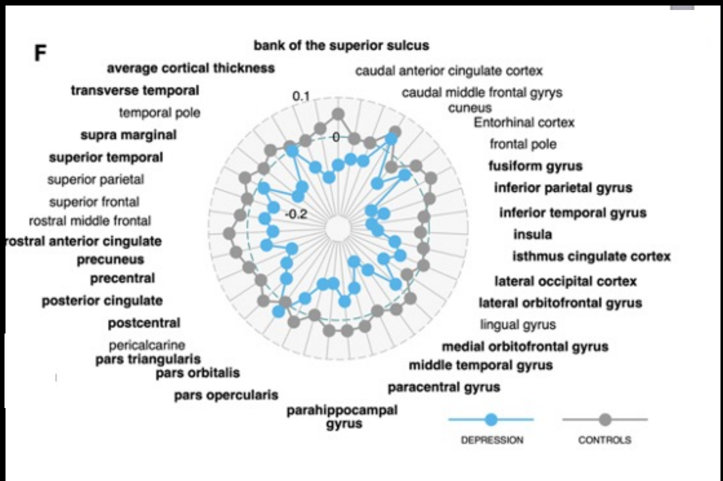
I.

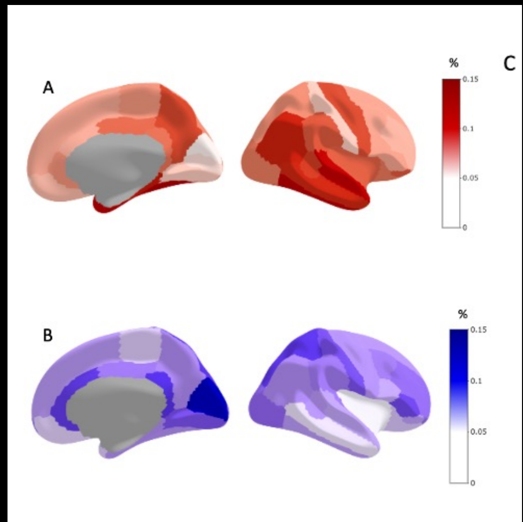
Parse heterogeneity

2.

Brain-behavior mappings

Extreme negative deviations





Extreme positive deviations

- earlier onset of depression
- more severe depressive symptoms
- higher BMI

- being remitted
- not taking antidepressants
- less severe symptoms

Bayer, J. M. M. et al. Dissecting heterogeneity in cortical thickness in major depressive disorder: a large scale normative modelling study. (2025) submitted to JAMA Network Open

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset

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

Bold Signal

Covariates (age, sex, Task parameters)

Model parameters

Residuals

Voxels →

Subjects →

Z-statistic image (Y)

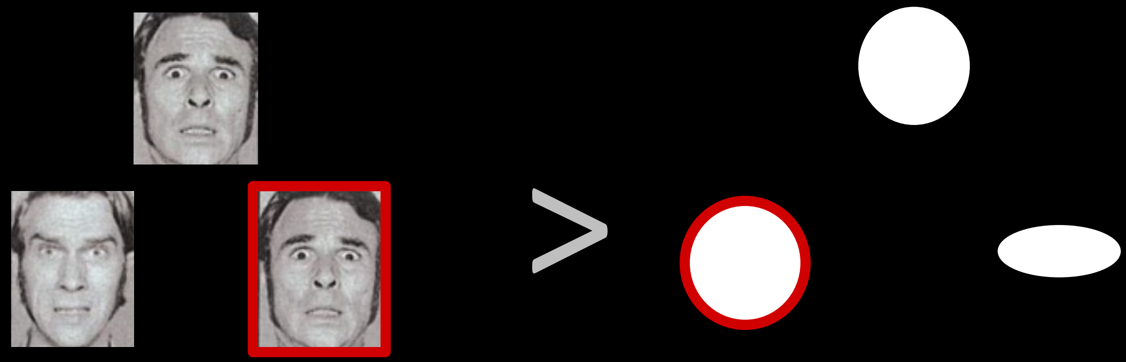
Predictors →

Subjects →

Age	Site
Sex	Scanner strength
	TR
Task length	Multiband-factor
# of target blocks	# of volumes acquired
Instructions	

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset



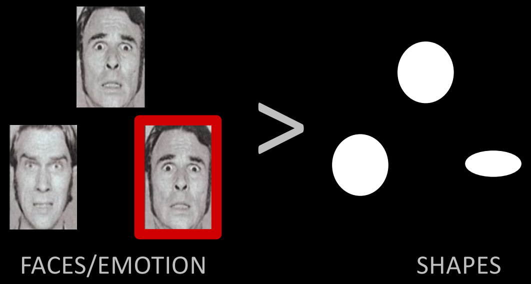
MATCH FACES/EMOTION

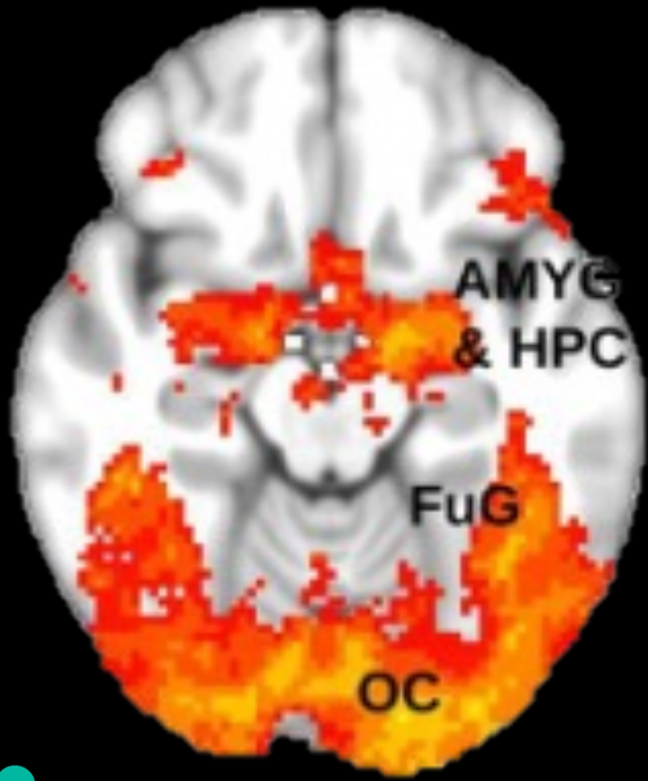
MATCH SHAPES

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

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset





Unaffected controls

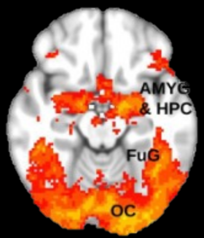
Viering et al (2021) doi: 10.1007/s00787-021-01809-3

APPLICATIONS: Functional Imaging

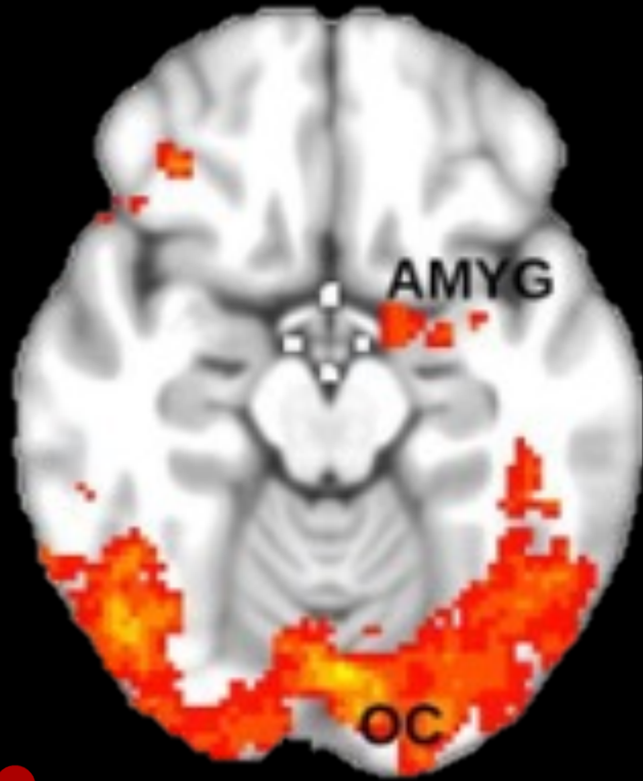
- Training and evaluating the normative model on local dataset

FACES/EMOTION

SHAPES



Unaffected controls

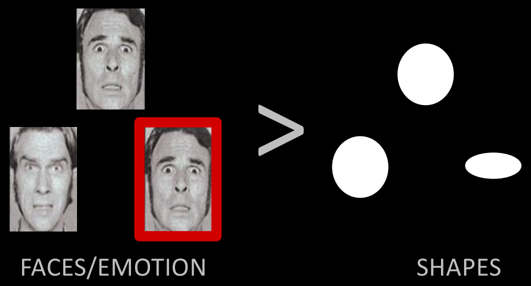


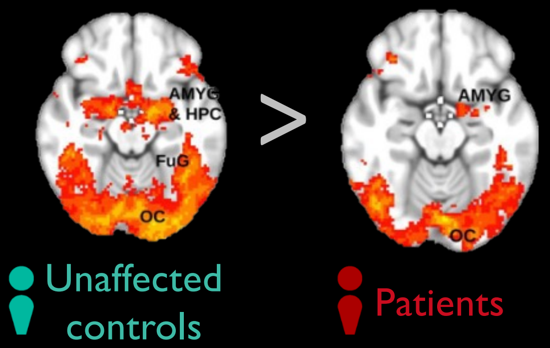
Patients

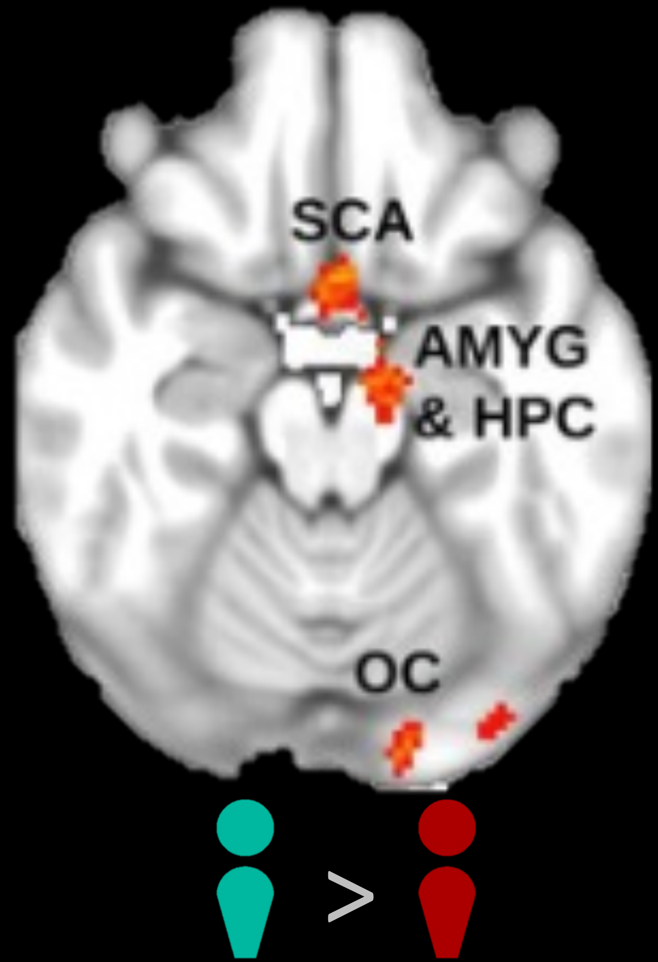
Viering et al (2021) doi: 10.1007/s00787-021-01809-3

APPLICATIONS: Functional Imaging

- Training and evaluating the normative model on local dataset





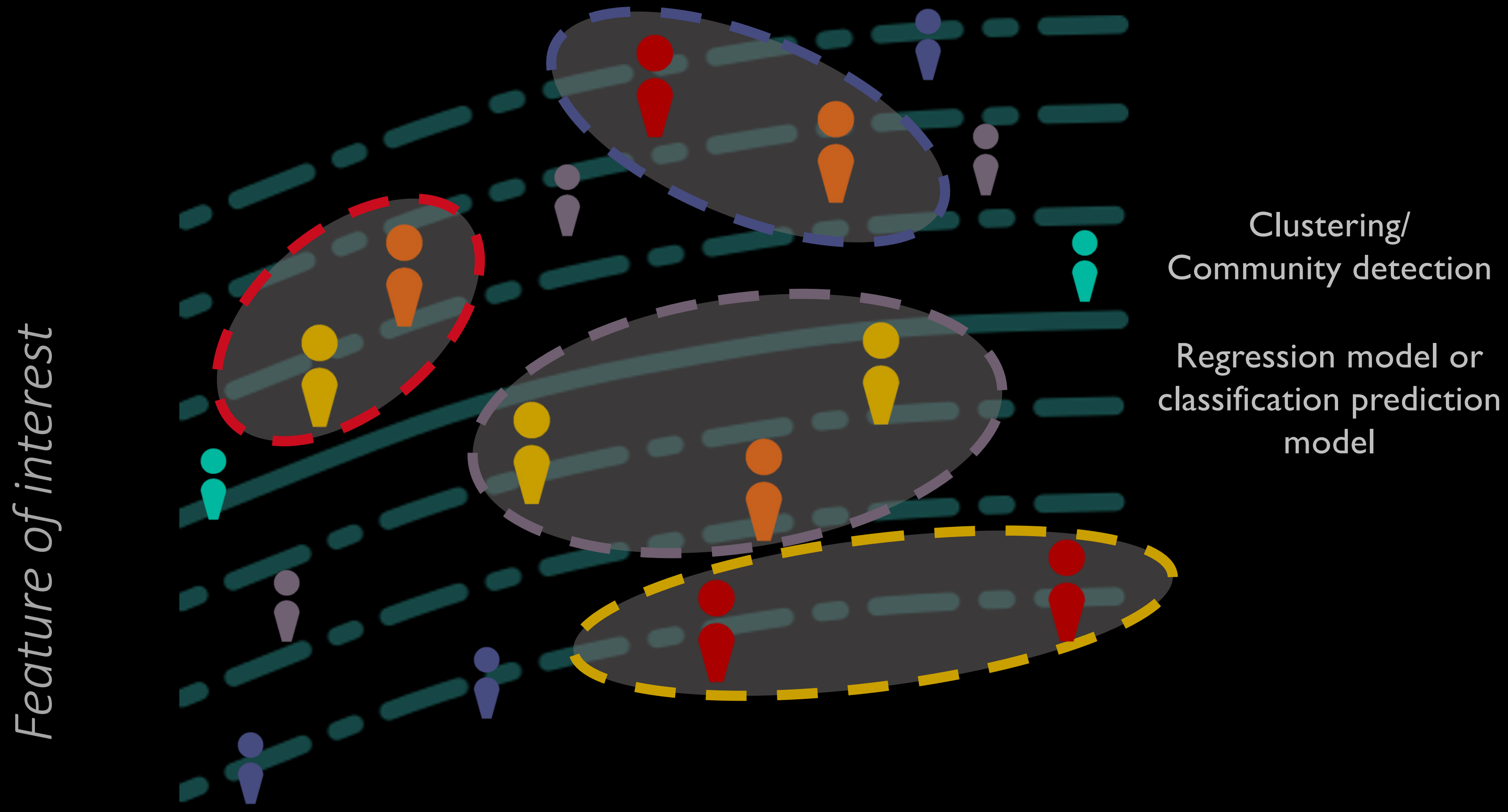


Viering et al (2021) doi: 10.1007/s00787-021-01809-3

HETEROGENEITY NORMATIVE MODELLING

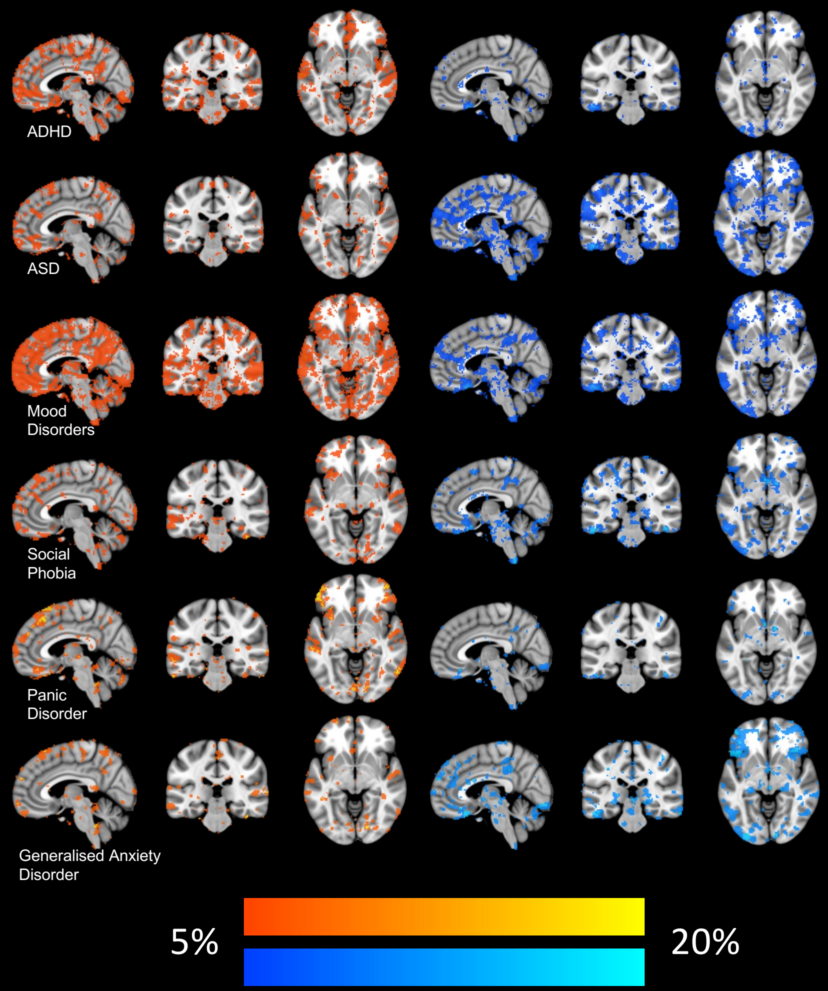
NEUROBIOLOGICAL SUBTYPING

APPLICATIONS



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.

APPLICATIONS & DEVELOPMENT

Training your own models

Applying pre-trained models

Methodology

APPLICATIONS & DEVELOPMENT

Training your own models

Applying pre-trained models

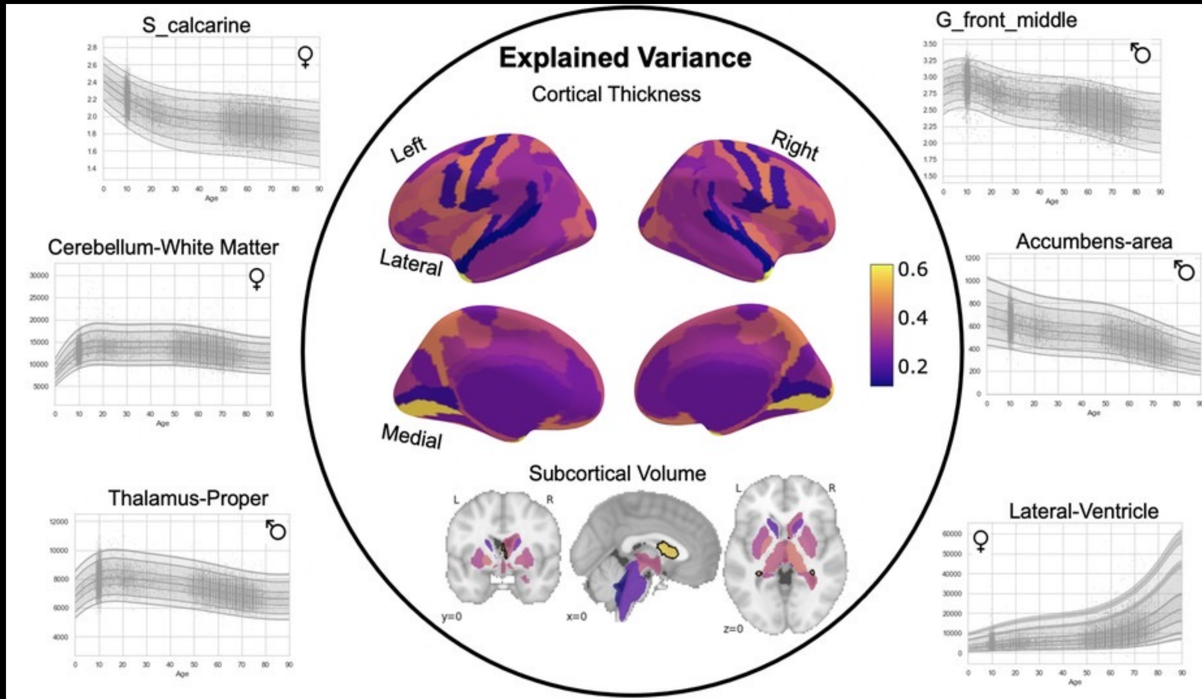
Methodology

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



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

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

USE PRETRAINED MODELS

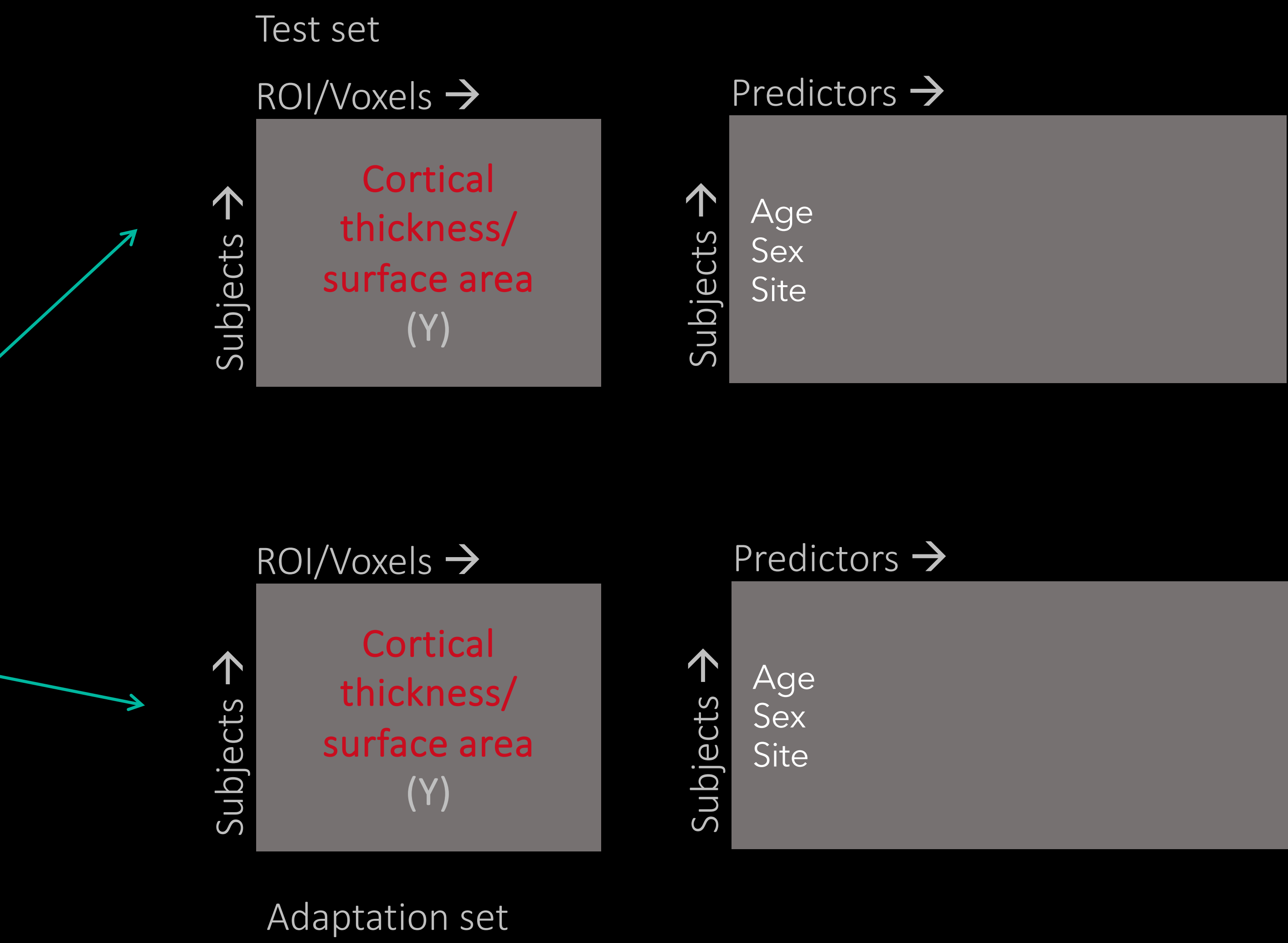
- Using pre-trained models, evaluating on local datasets

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

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

NORMATIVE MODELS

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

Brain (BOLD signal in voxel, ROI)

Covariates (age, sex, task parameters)

Model parameters

Residuals

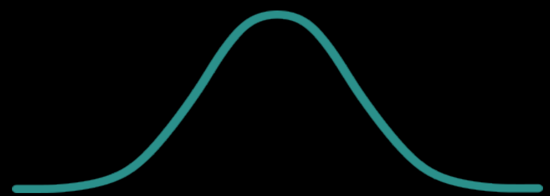
Gaussian process regression

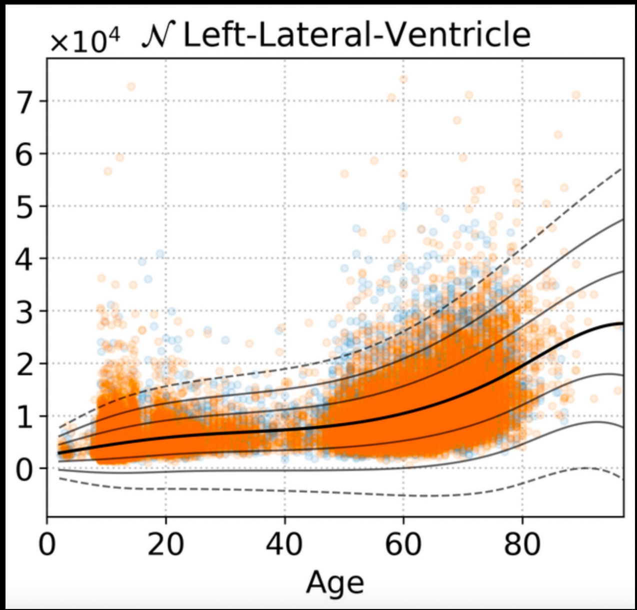
Warped Bayesian linear regression

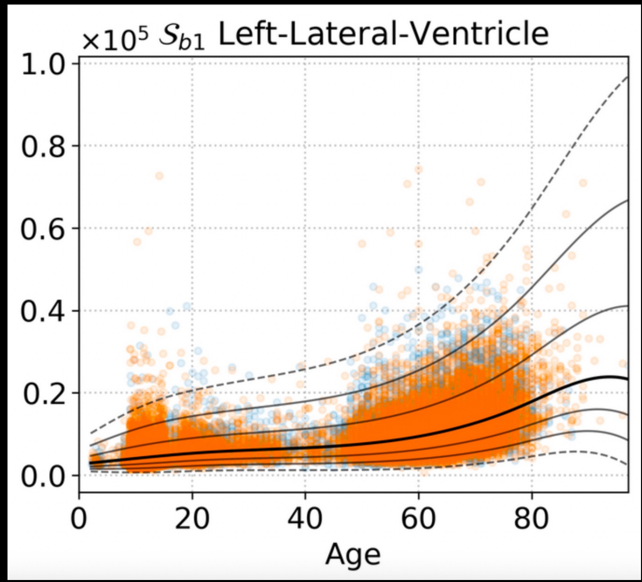
Generalized additive models of location scale and shape

Hierarchical Bayesian regression

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



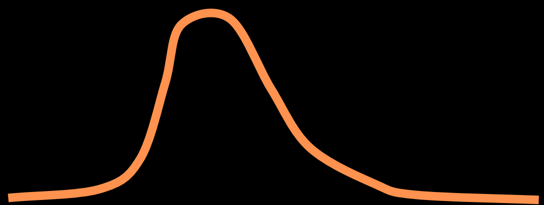


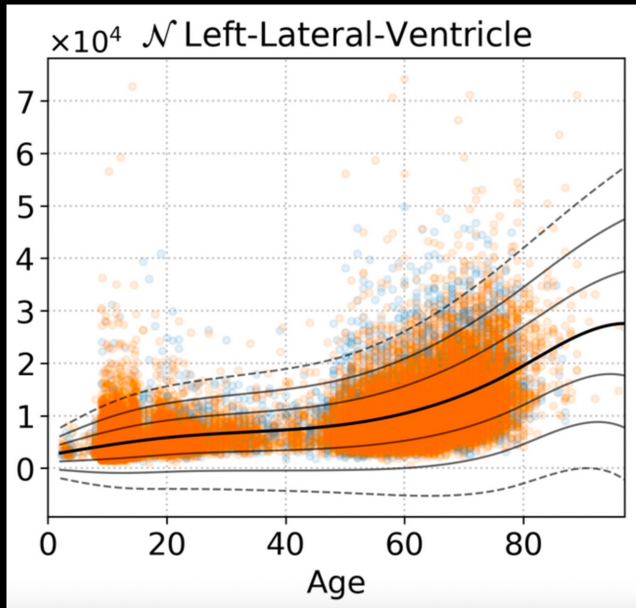


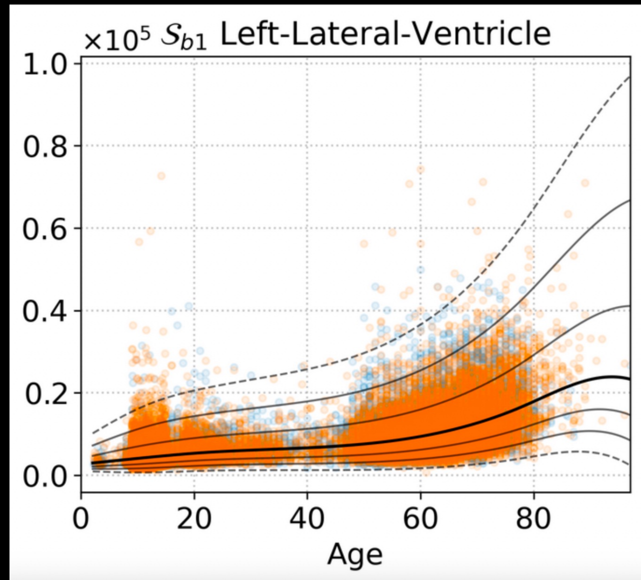
de Boer, A. A. A. Bayer, J. M. M. *et al.* (2024) **Non-Gaussian normative modelling with hierarchical Bayesian regression**. *Imaging Neuroscience* **2**, 1–36.

APPLICATIONS

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

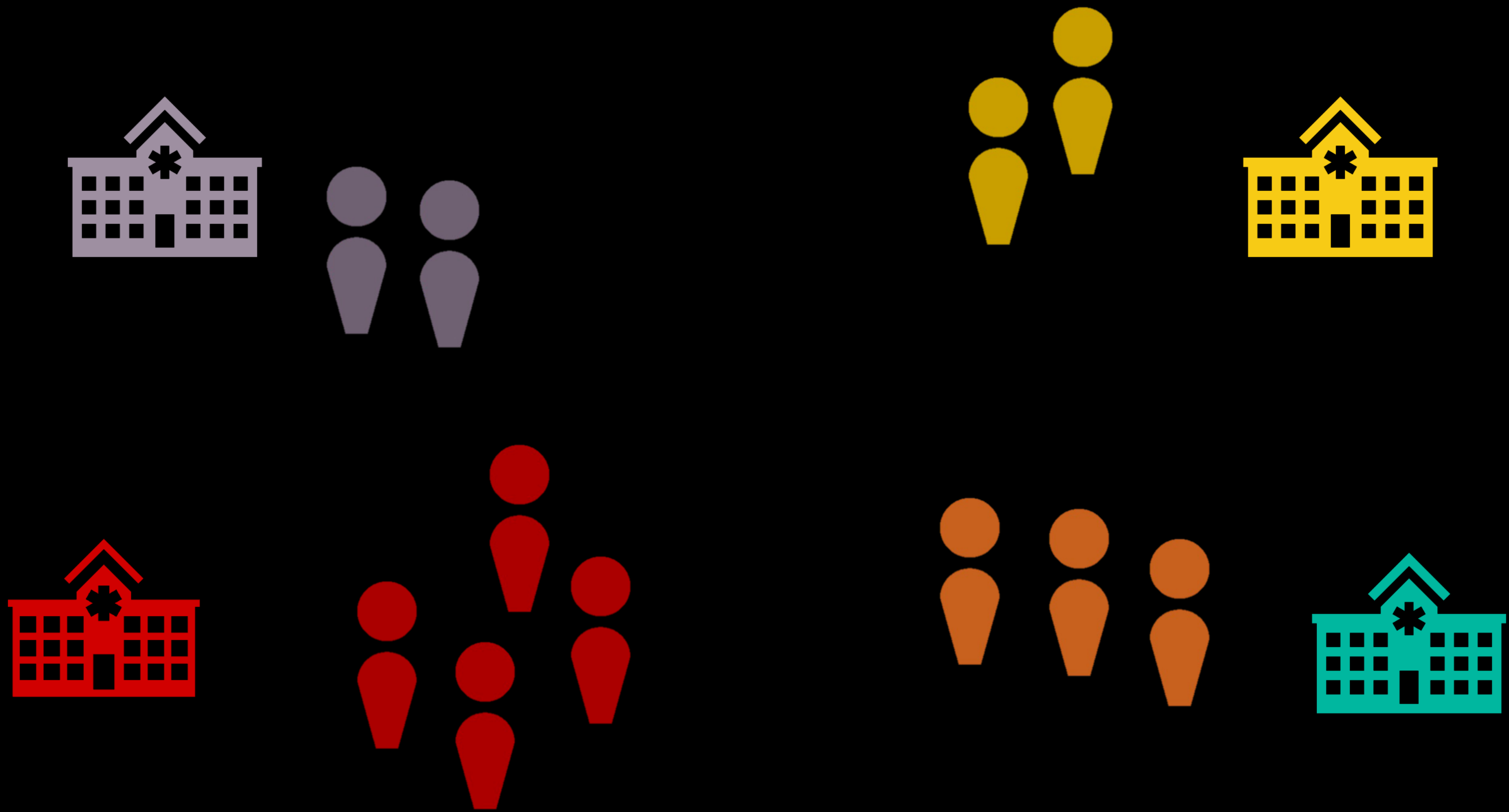






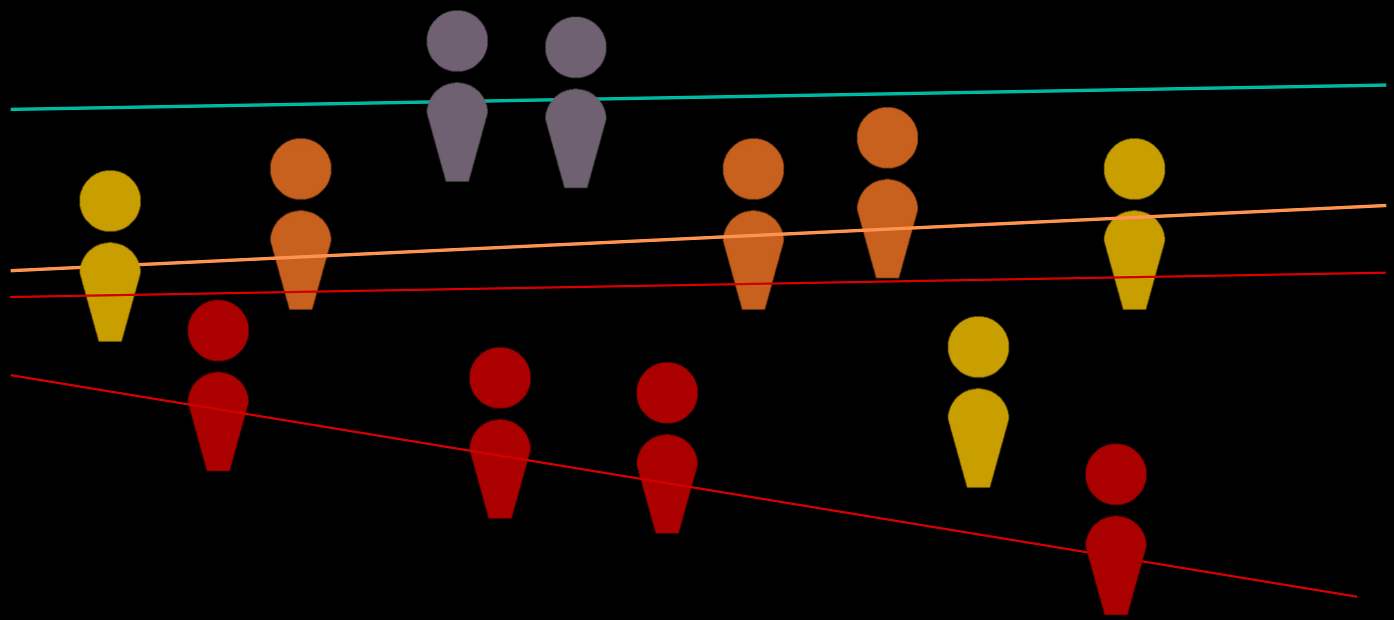
de Boer, A. A. A. Bayer, J. M. M. *et al.* (2024) **Non-Gaussian normative modelling with hierarchical Bayesian regression**. *Imaging Neuroscience* **2**, 1–36.

APPLICATIONS



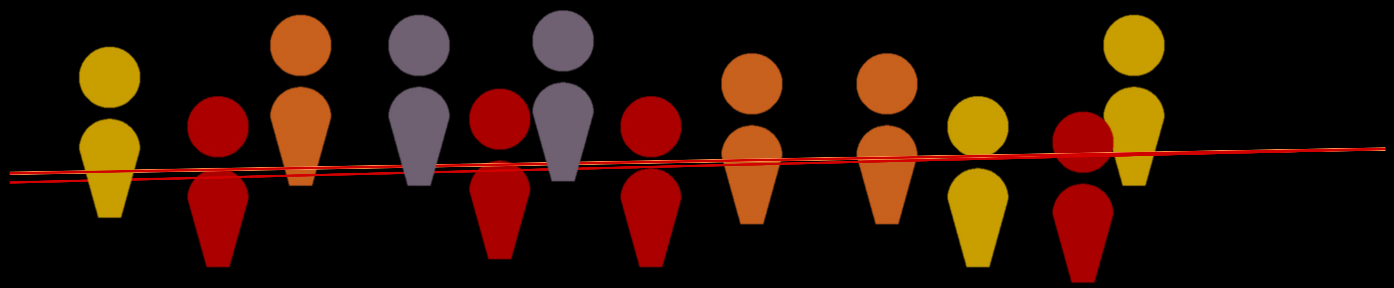
Bayer, J. M. M. *et al.* Site effects how-to and when: An overview of retrospective techniques to accommodate site effects in multi-site neuroimaging analyses. *Front. Neurol.* **13**, (2022).

APPLICATIONS



Bayer, J. M. M. *et al.* Site effects how-to and when: An overview of retrospective techniques to accommodate site effects in multi-site neuroimaging analyses. *Front. Neurol.* **13**, (2022).

APPLICATIONS

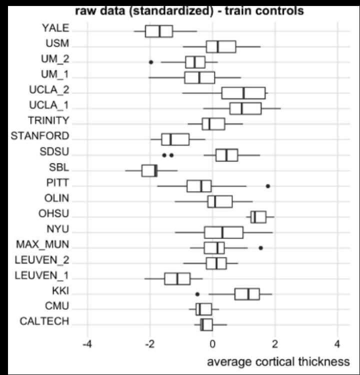


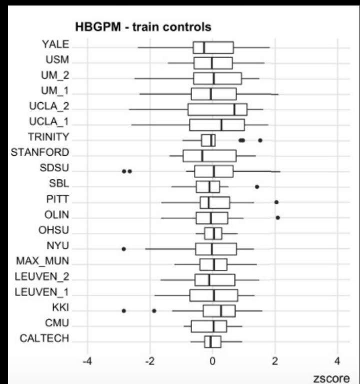
Bayer, J. M. M. *et al.* Site effects how-to and when: An overview of retrospective techniques to accommodate site effects in multi-site neuroimaging analyses. *Front. Neurol.* **13**, (2022).

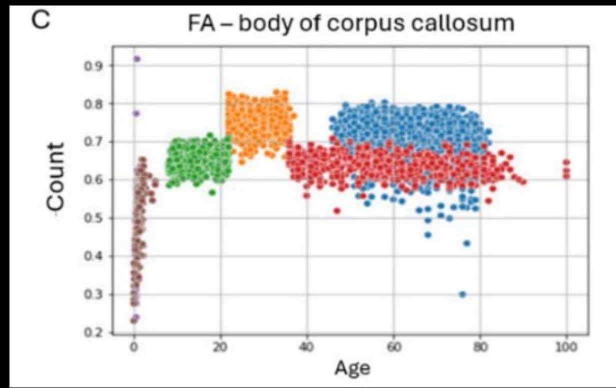
APPLICATIONS

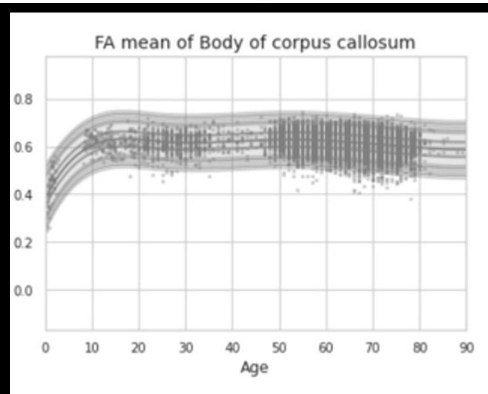
- Including site as batch effect into normative model

- Inverting z-scores between sites







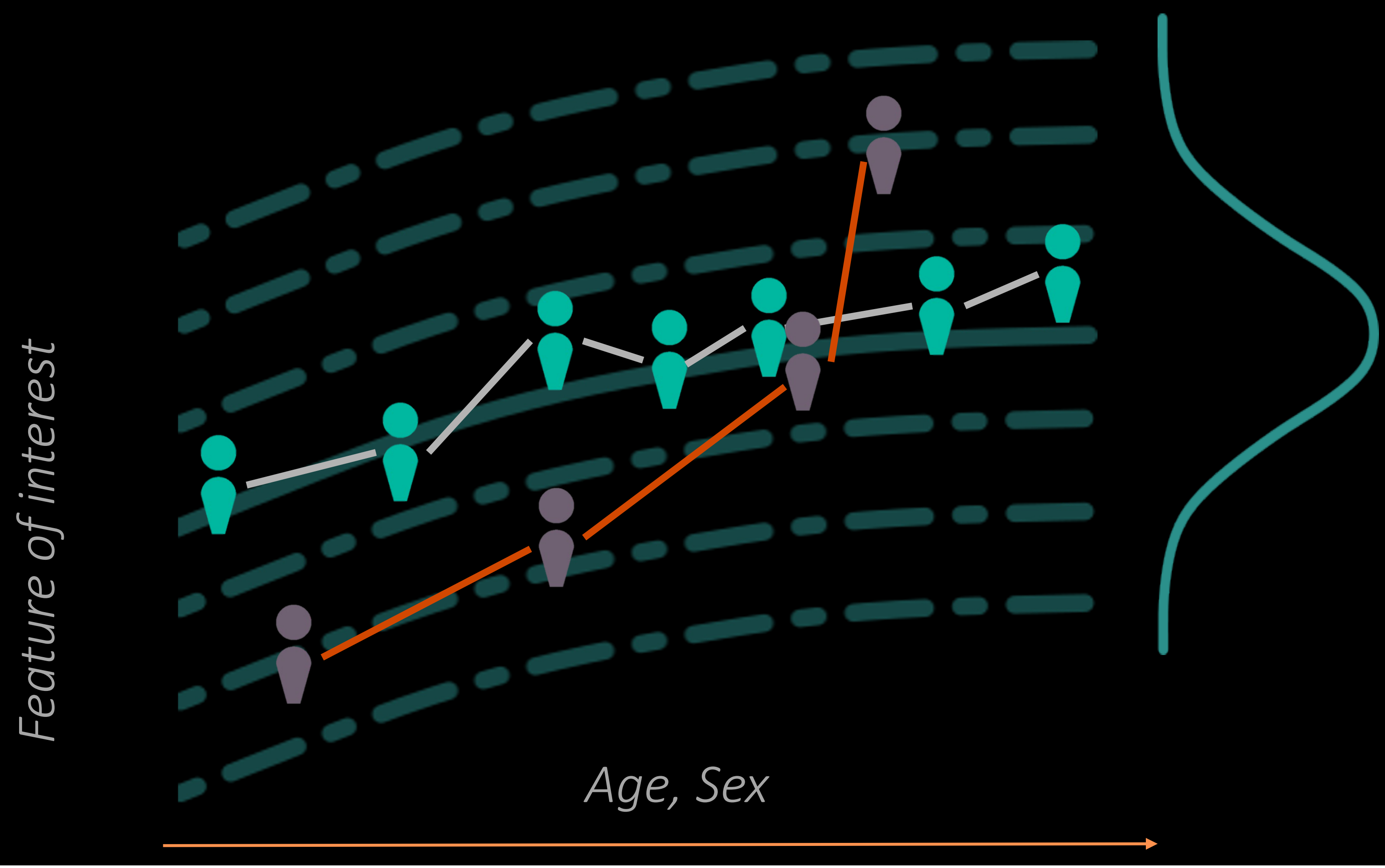


Bayer, J. M. M. *et al.* Accommodating site variation in neuroimaging data using normative and hierarchical Bayesian models. *Neuroimage* **264**, 119699 (2022).

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

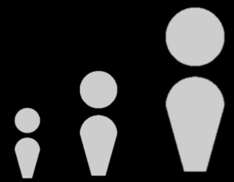
LONGITUDINAL

NORMATIVE MODELLING



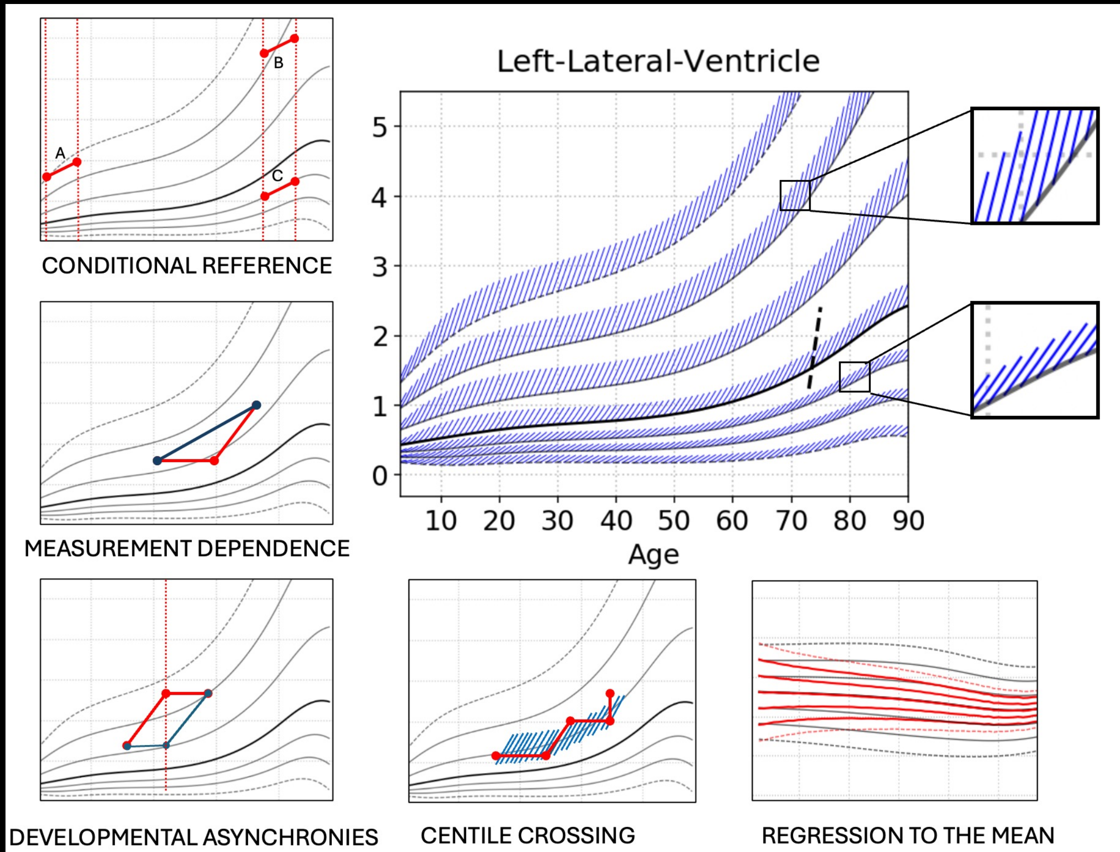
OTHER

APPLICATIONS



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.

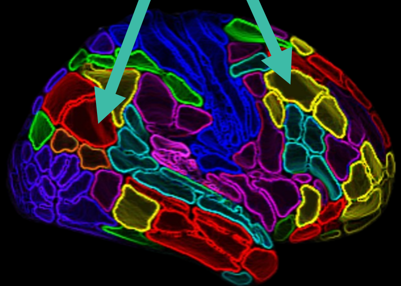
ESTIMATING LIFESPAN NORMATIVE MODELS

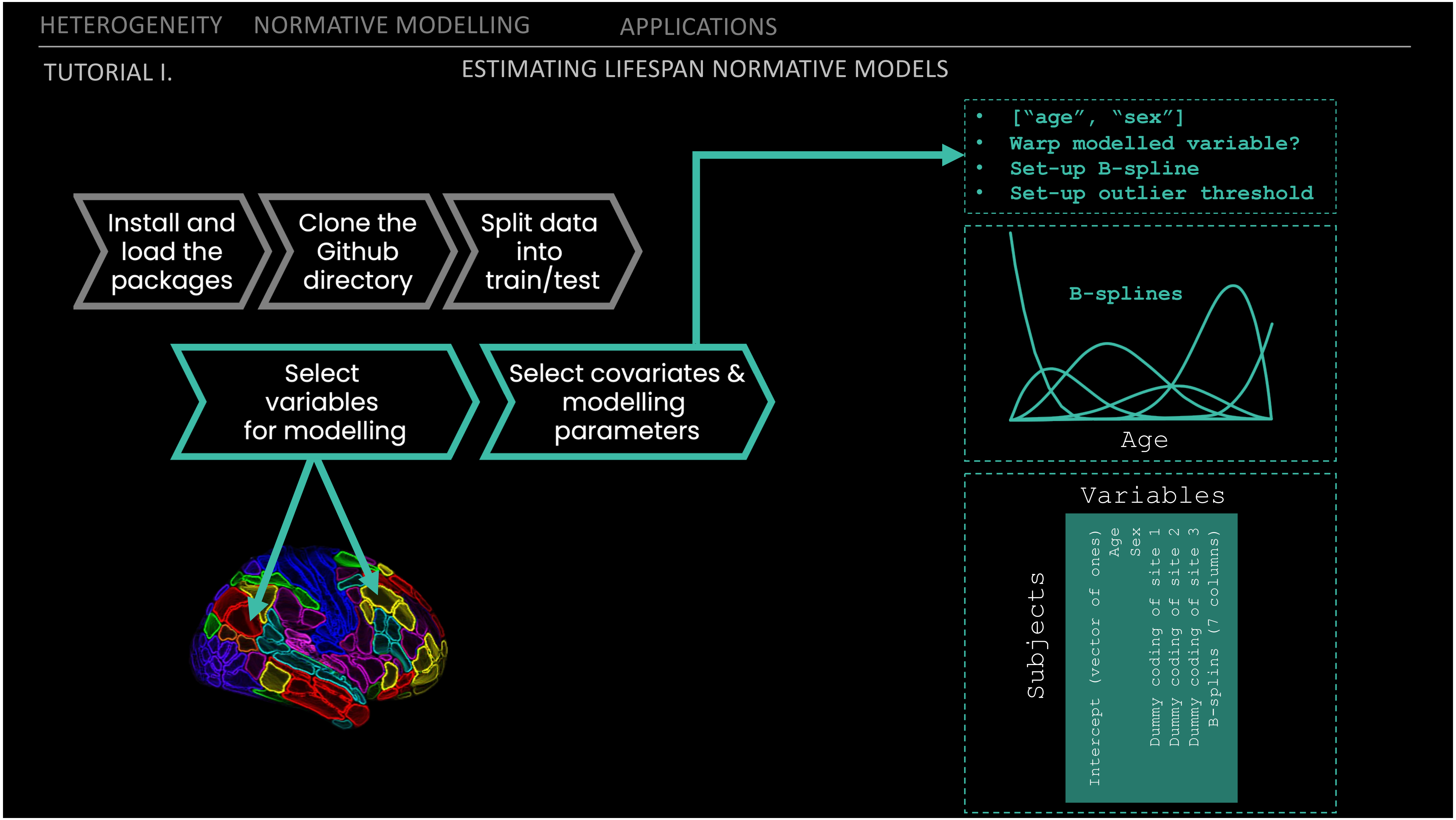
TUTORIAL I.

ESTIMATING LIFESPAN NORMATIVE MODELS

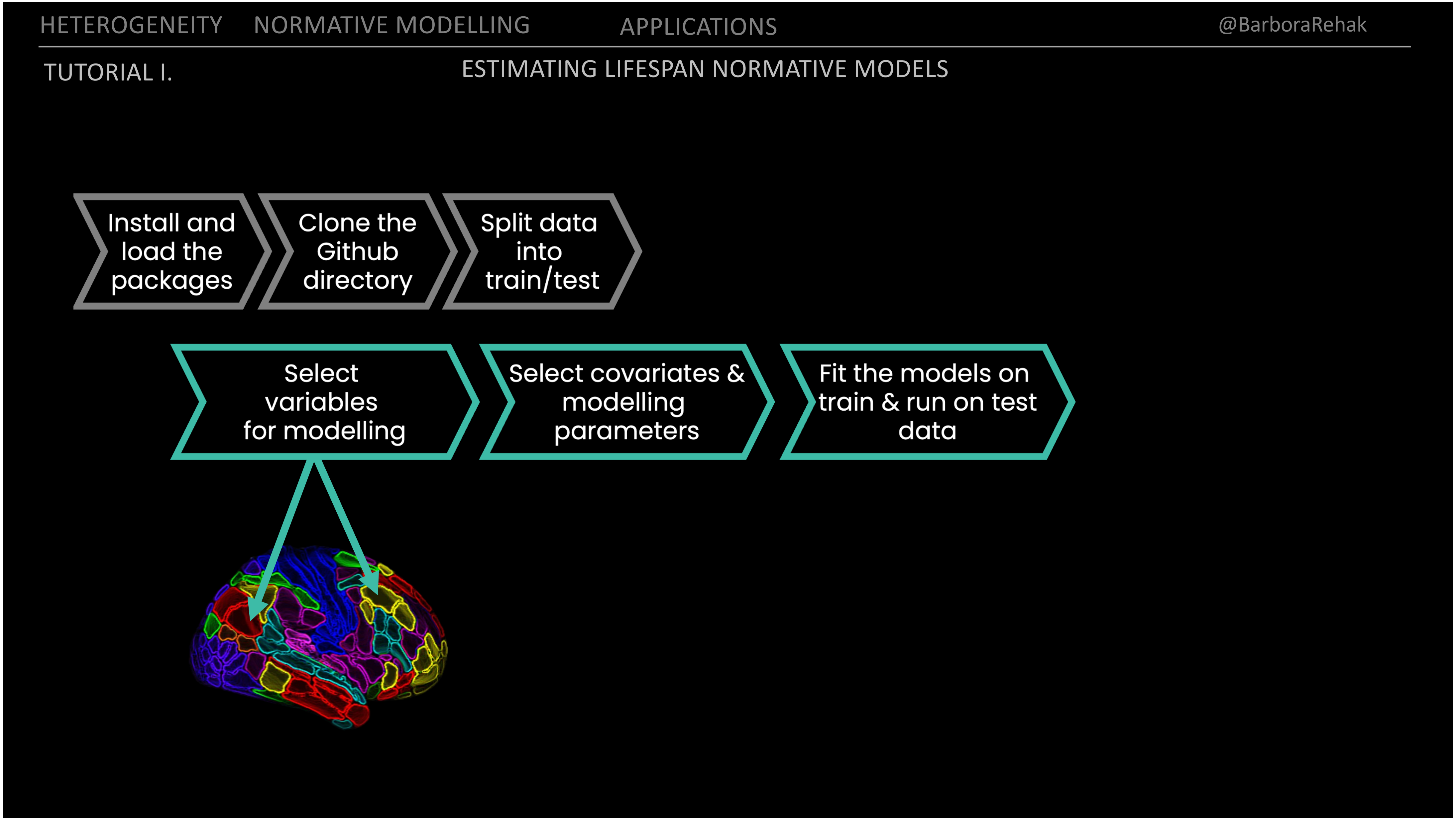
Install and load the packages → Clone the Github directory → Split data into train/test

Select variables for modelling → Select covariates & modelling parameters





HETEROGENEITY
NORMATIVE MODELLING
APPLICATIONS
TUTORIAL I.
ESTIMATING LIFESPAN NORMATIVE MODELS
Install and load the packages
Clone the Github directory
Split data into train/test
Select variables for modelling
Select covariates & modelling parameters
["age", "sex"]
Warp modelled variable?
Set-up B-spline
Set-up outlier threshold
B-splines
Age
Variables
Subjects
Intercept (vector of ones)
Age
Sex
Dummy coding of site 1
Dummy coding of site 2
Dummy coding of site 3
B-splins (7 columns)


HETEROGENEITY
NORMATIVE MODELLING
APPLICATIONS
@BarboraRehak
TUTORIAL I.
ESTIMATING LIFESPAN NORMATIVE MODELS
Install and load the packages
Clone the Github directory
Split data into train/test
Select variables for modelling
Select covariates & modelling parameters
Fit the models on train & run on test data

TUTORIAL I.

ESTIMATING LIFESPAN NORMATIVE MODELS

Install and load the packages

Clone the Github directory

Split data into train/test

Select variables for modelling

Select covariates & modelling parameters

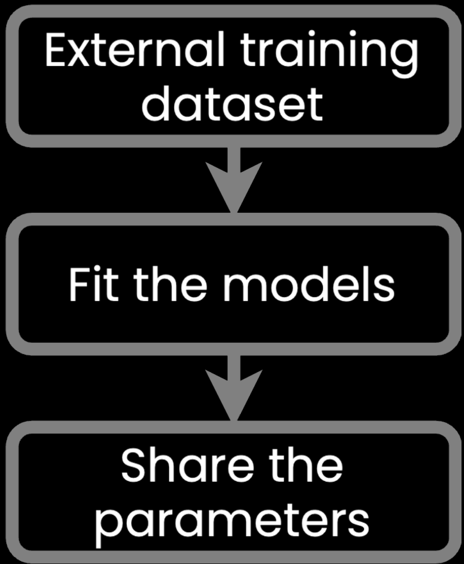
Fit the models on train & run on test data

- Negative log likelihood
- Explained variance
- Mean standardized log loss
- Bayesian information criteria
- Skew
- Kurtosis

Check the fit of the models

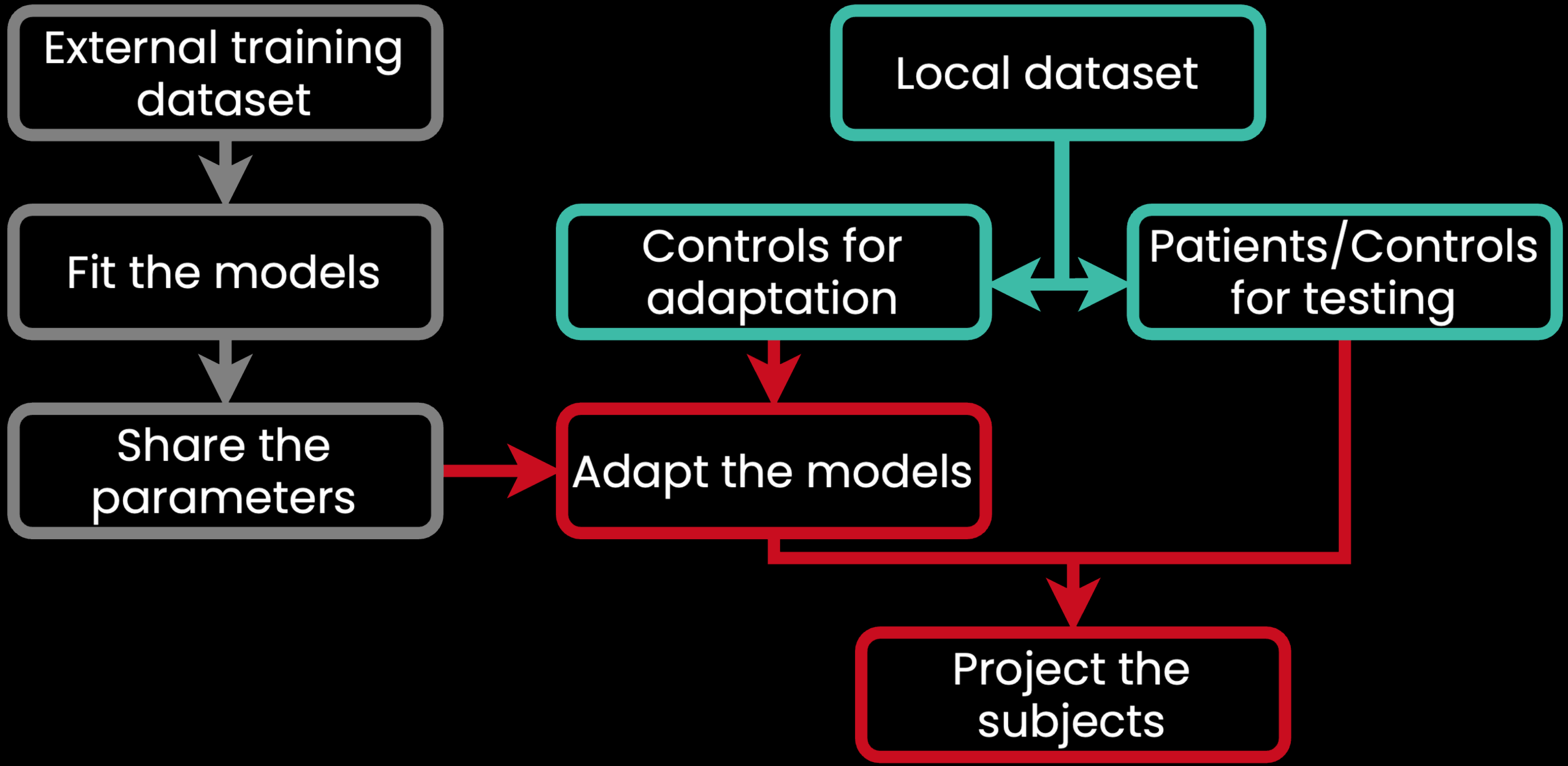
TUTORIAL II.

APPLYING PRE-TRAINED NORMATIVE MODELS



TUTORIAL II.

APPLYING PRE-TRAINED NORMATIVE MODELS



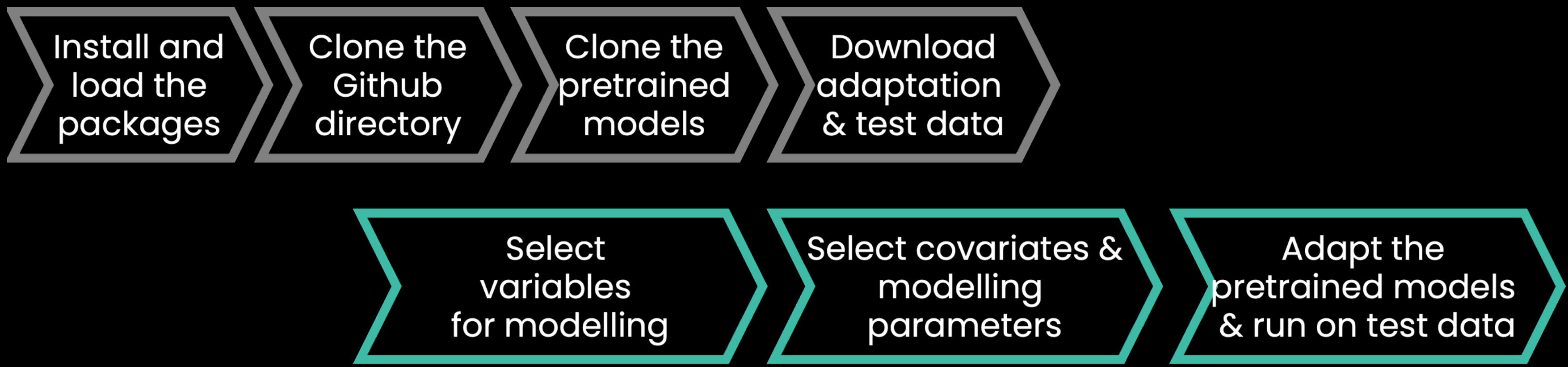
TUTORIAL II.

APPLYING PRE-TRAINED NORMATIVE MODELS

Install and load the packages → Clone the Github directory → Clone the pretrained models → Download adaptation & test data

TUTORIAL II.

APPLYING PRE-TRAINED NORMATIVE MODELS



TUTORIAL II.

APPLYING PRE-TRAINED NORMATIVE MODELS

Install and load the packages

Clone the Github directory

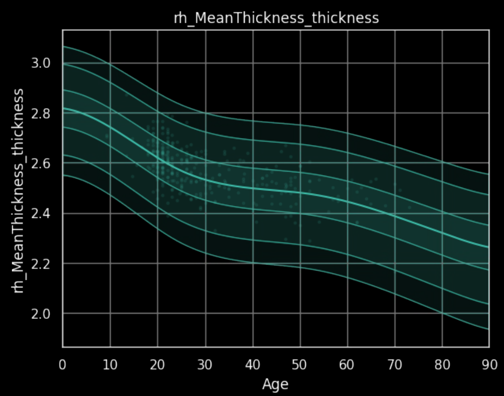
Clone the pretrained models

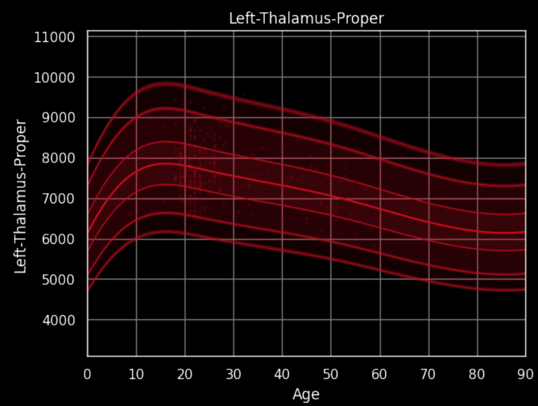
Download adaptation & test data

Select variables for modelling

Select covariates & modelling parameters

Adapt the pretrained models & run on test data





Visualize

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