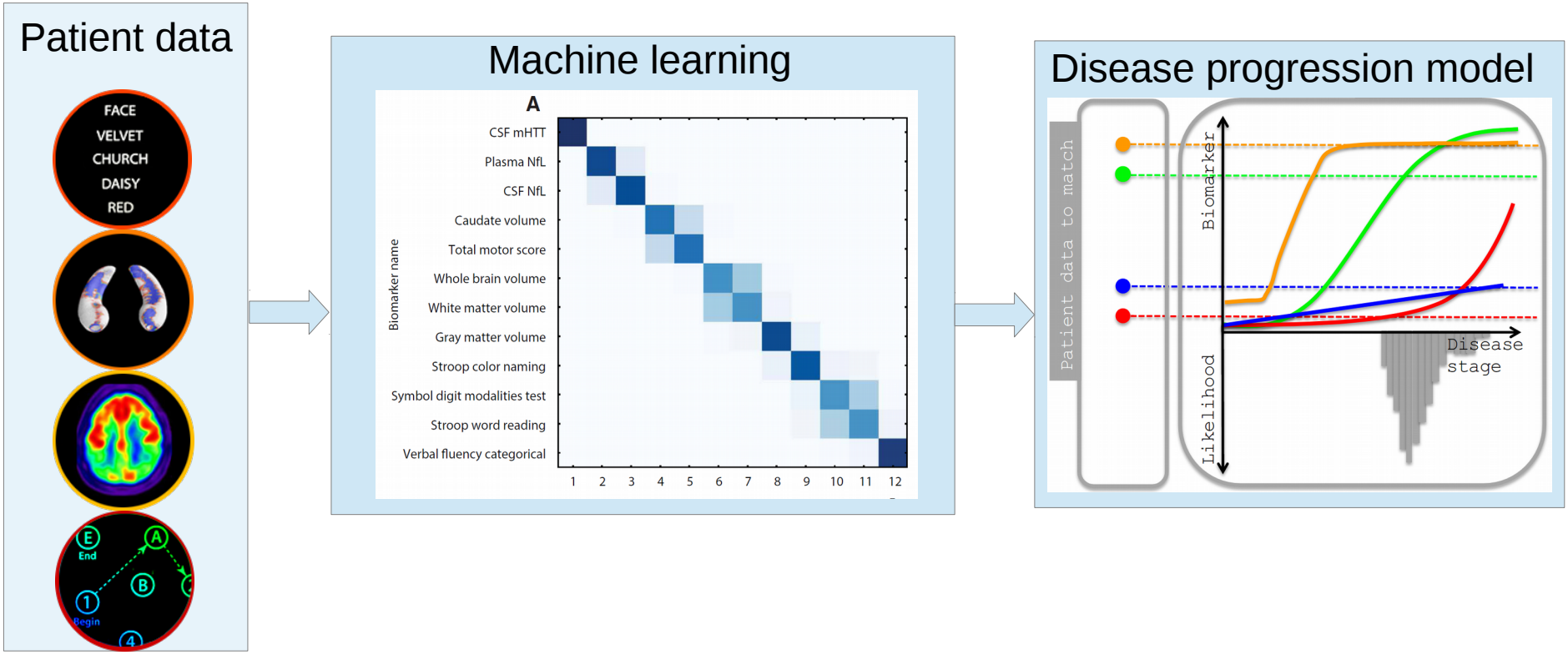
Computational models for clinical trial design in Huntington's disease

Peter Wijeratne



Acknowledgements

UCL CMIC

Daniel Alexander
Leon Aksman
Maura Bellio
Arman Eshaghi
Neil Oxtoby
Alexandra Young

UCL HDC

Sarah Tabrizi
Rachael Scahill
Sarah Gregory
Eileanoir Johnson
Ed Wild
Lauren Byrne

CHDI

Cristina Sampaio
Amrita Mohan
John Warner
Dorian Pustina
Alexandra Shechtel

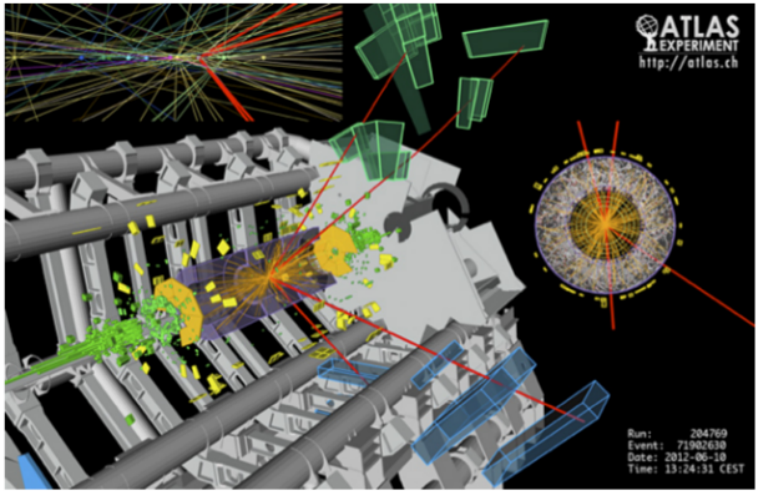
And all the participants of the PREDICT, TRACK and IMAGE-HD studies

MRC London Institute of Medical Sciences

CHDI FOUNDATION

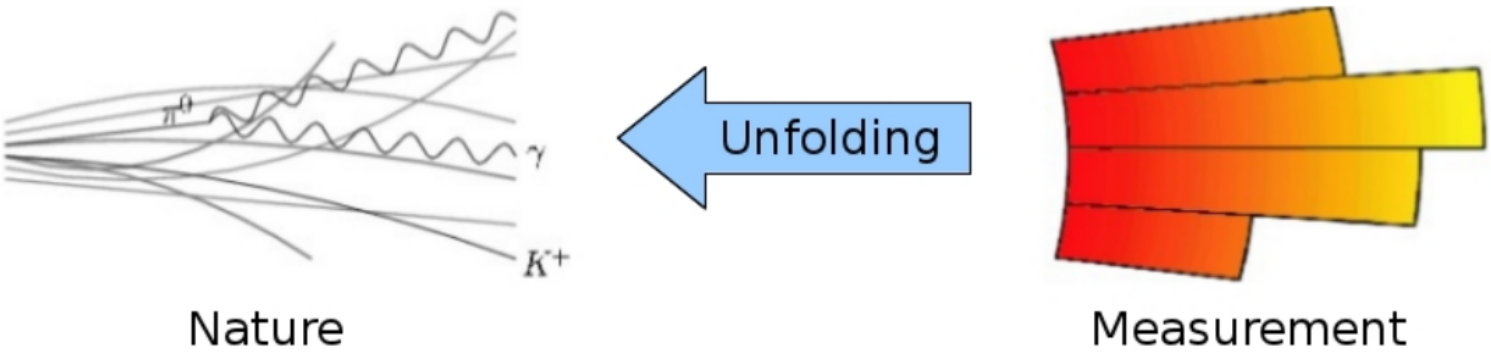
EPSRC Engineering and Physical Sciences Research Council

Background: my former life as a particle physicist



(a) A Higgs-like boson

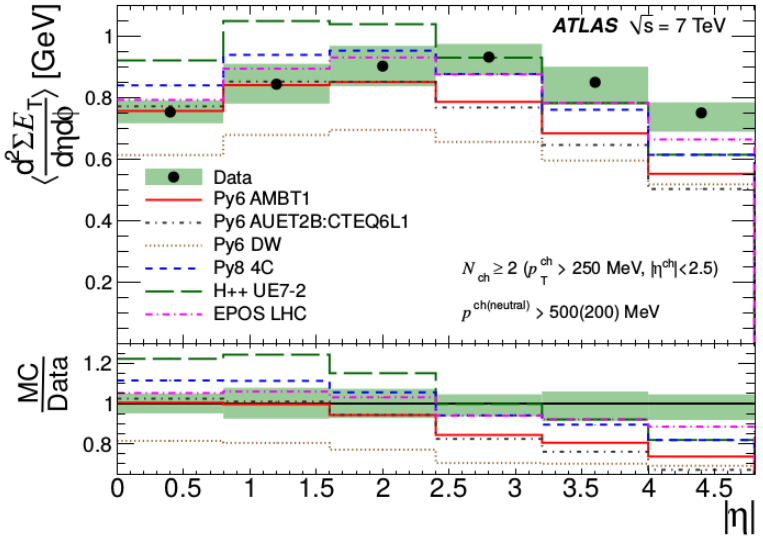
(b) CERN sheep outside ATLAS



$$n(C_i^{data}) = \frac{1}{\epsilon_i} \sum_j P(T_i^{MC} \mid R_j^{MC}) n(R_j^{data})$$

- Real data are dependent on the detector used to measure them

Bring data back to their natural state by applying hypothesis-driven corrections derived from simulation



I saw this one day in 2013

I wanted to use physics to fight cancer

I asked about for potential opportunities

I got lucky and a postdoc came up at the Centre for Medical Image Computing

Centre for Medical Image Computing (CMIC)

Maths, physics and engineering scientists at the interface of basic and biomedical sciences

CMIC

Great Ormond Street Hospital

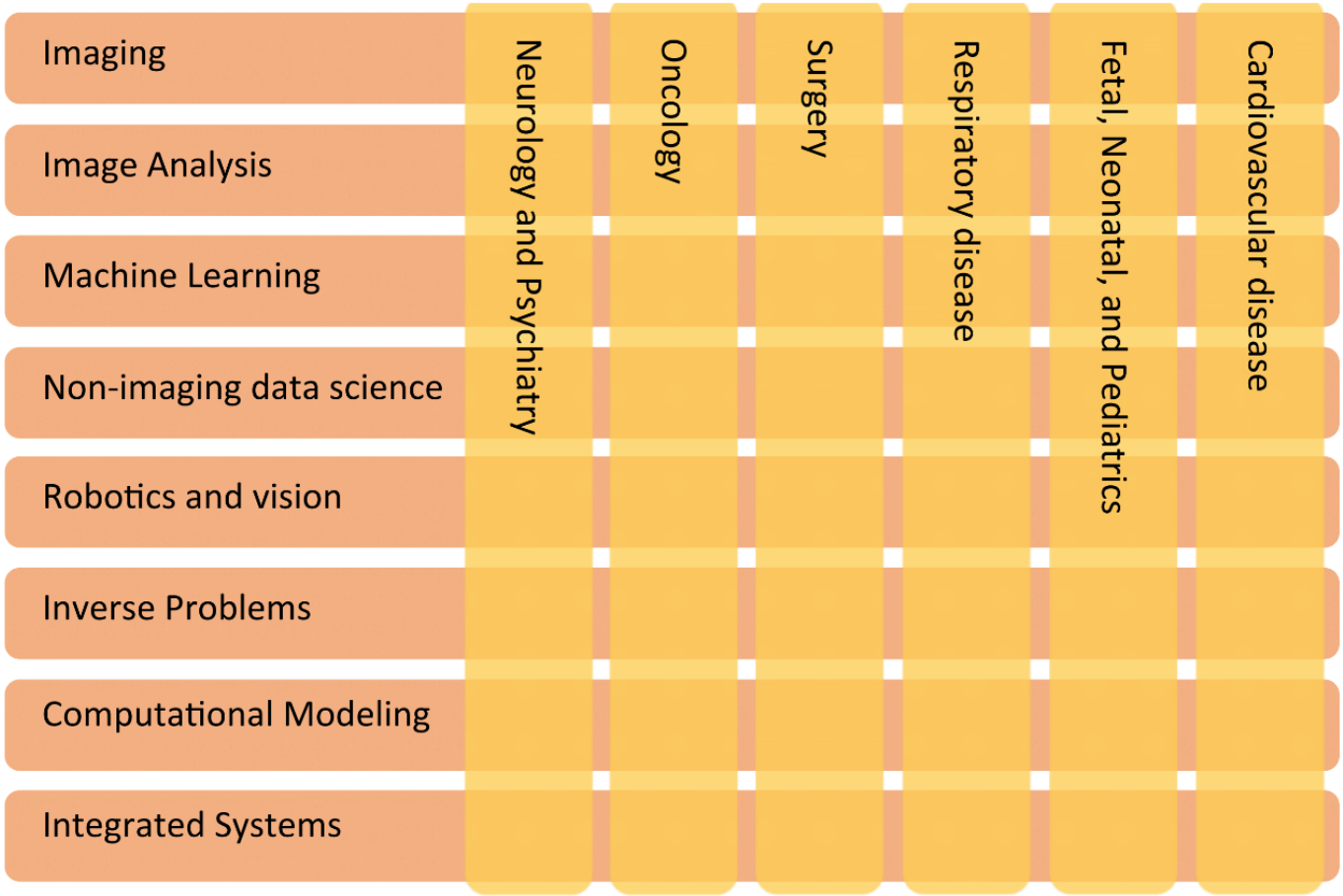
University College London Hospital

Moorfield's Eye Hospital

Royal Free Hospital

Royal National Orthopaedic Hospital

CMIC capability-application



Slight (3 year) diversion: biophysical modelling of drug delivery

The Chemical Basis of Morphogenesis

A. M. Turing

Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, Vol. 237, No. 641. (Aug. 14, 1952), pp. 37-72.

Computational Modeling

$$\frac{\partial u}{\partial t} = F(u,v) - d_u v + D_u \Delta u$$

$$\frac{\partial v}{\partial t} = G(u,v) - d_v v + D_v \Delta v$$

Rate of concentration change | Production | Degradation | Diffusion

Reaction

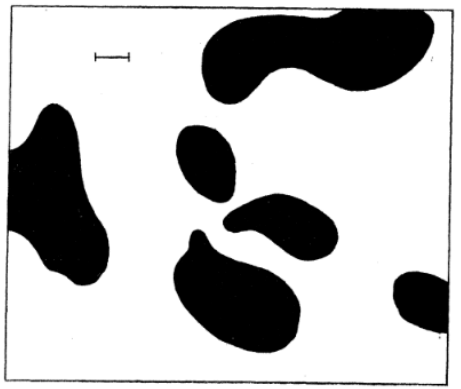
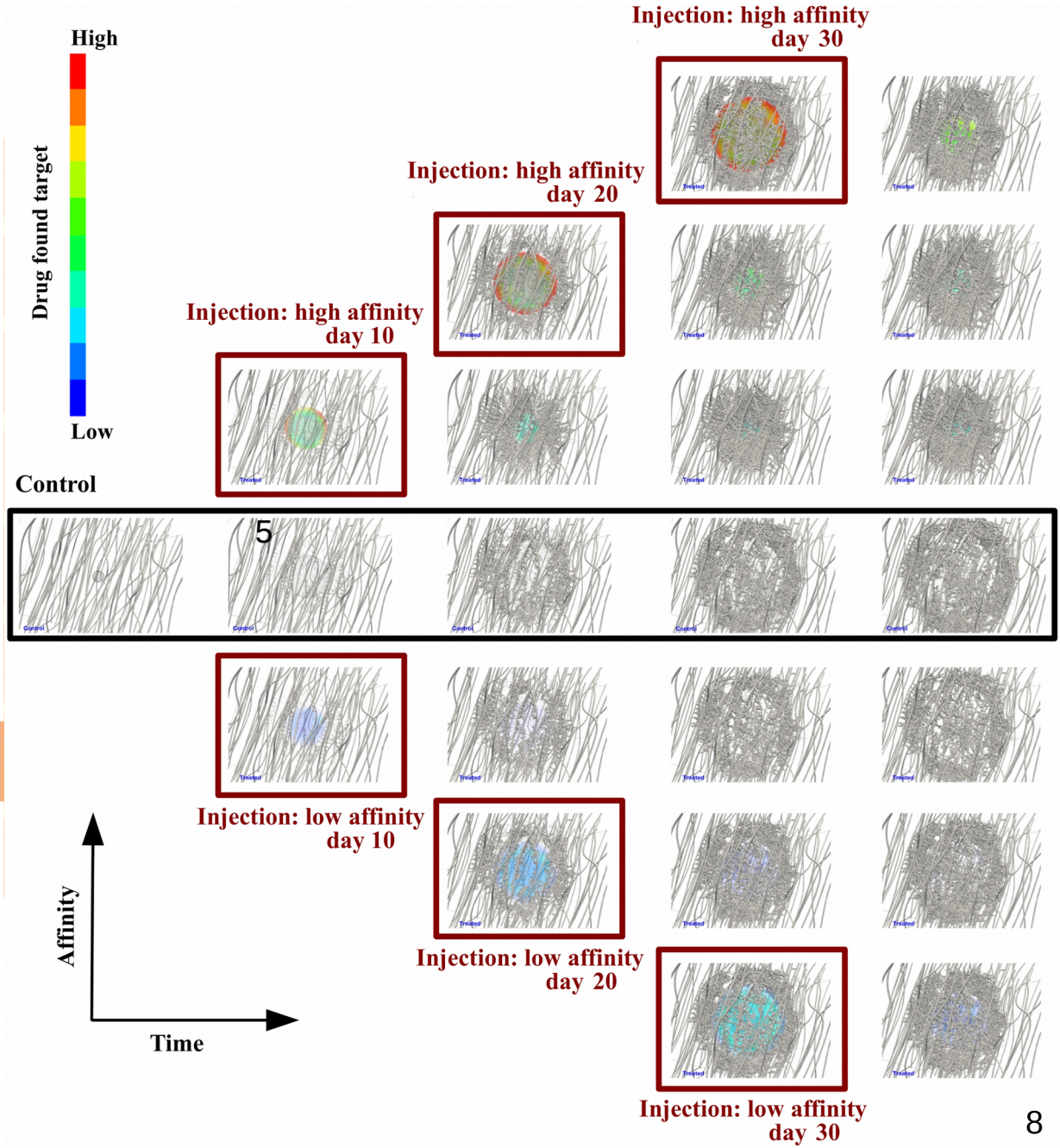


FIGURE 2. An example of a 'dappled' pattern as resulting from a type (*a*) morphogen system. A marker of unit length is shown. See text, §9, 11.

Slight (3 year) diversion: biophysical modelling of drug delivery

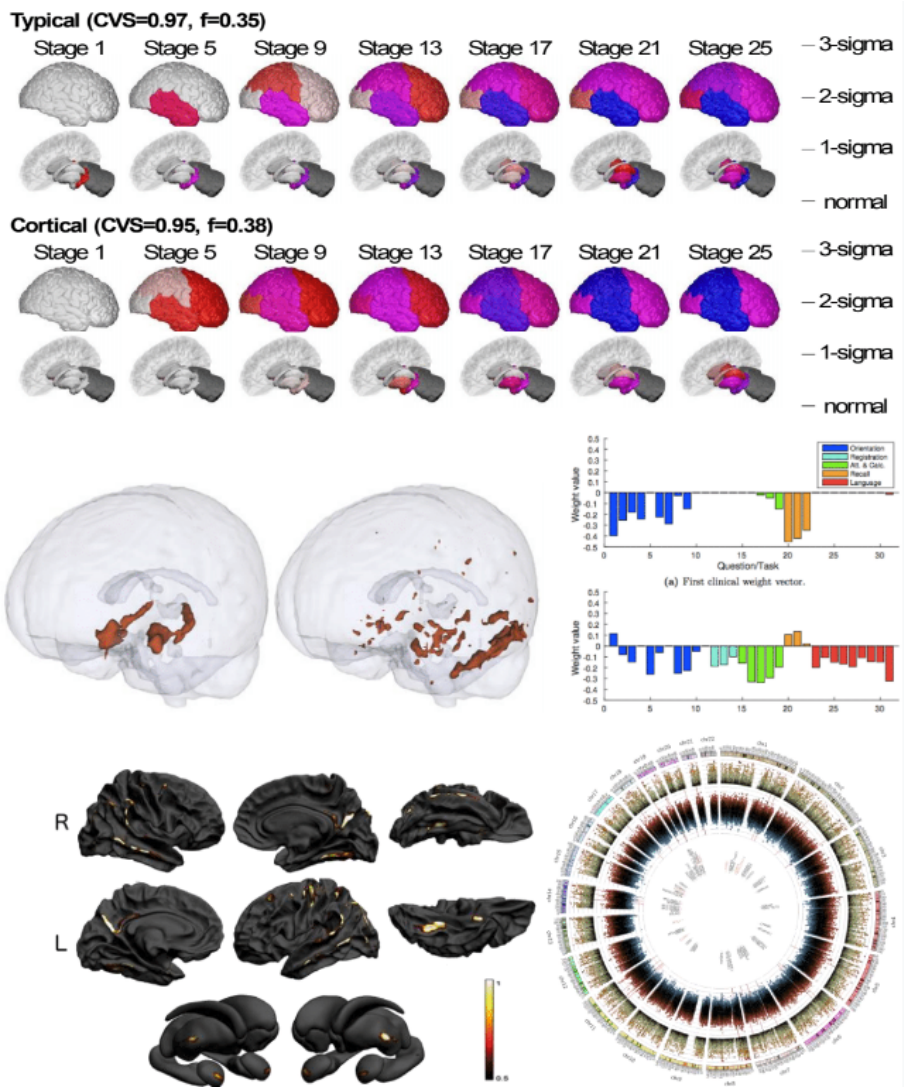
Computational Modeling



Vavourakis, Stylianopoulos, Wijeratne (2018) PLOS Comp Bio

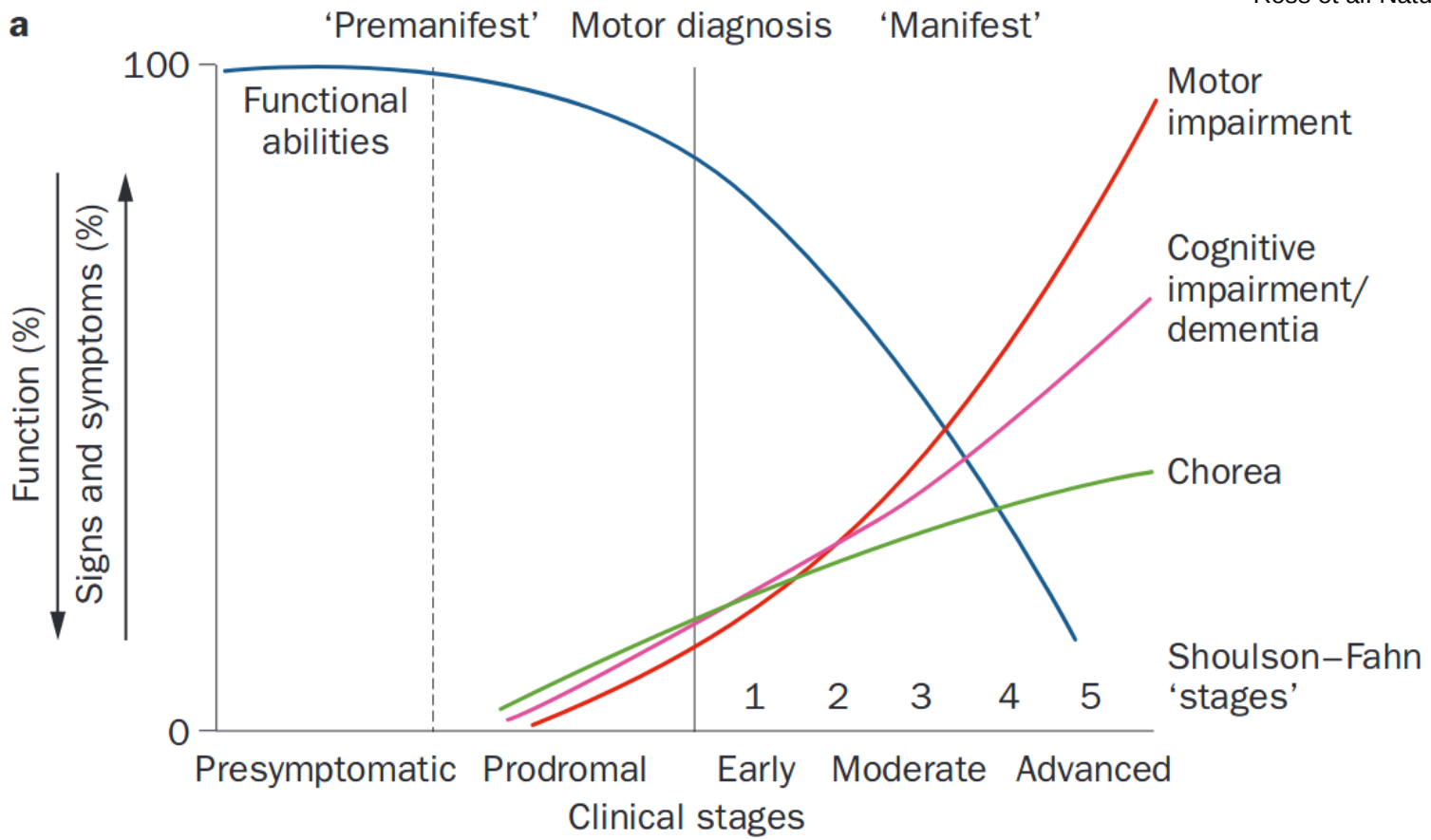
Current work: computational neurology

- Image Analysis
- Machine Learning
- Non-imaging data science
- Computational Modeling



Huntington's disease

Ross et al. Nature Reviews 2014

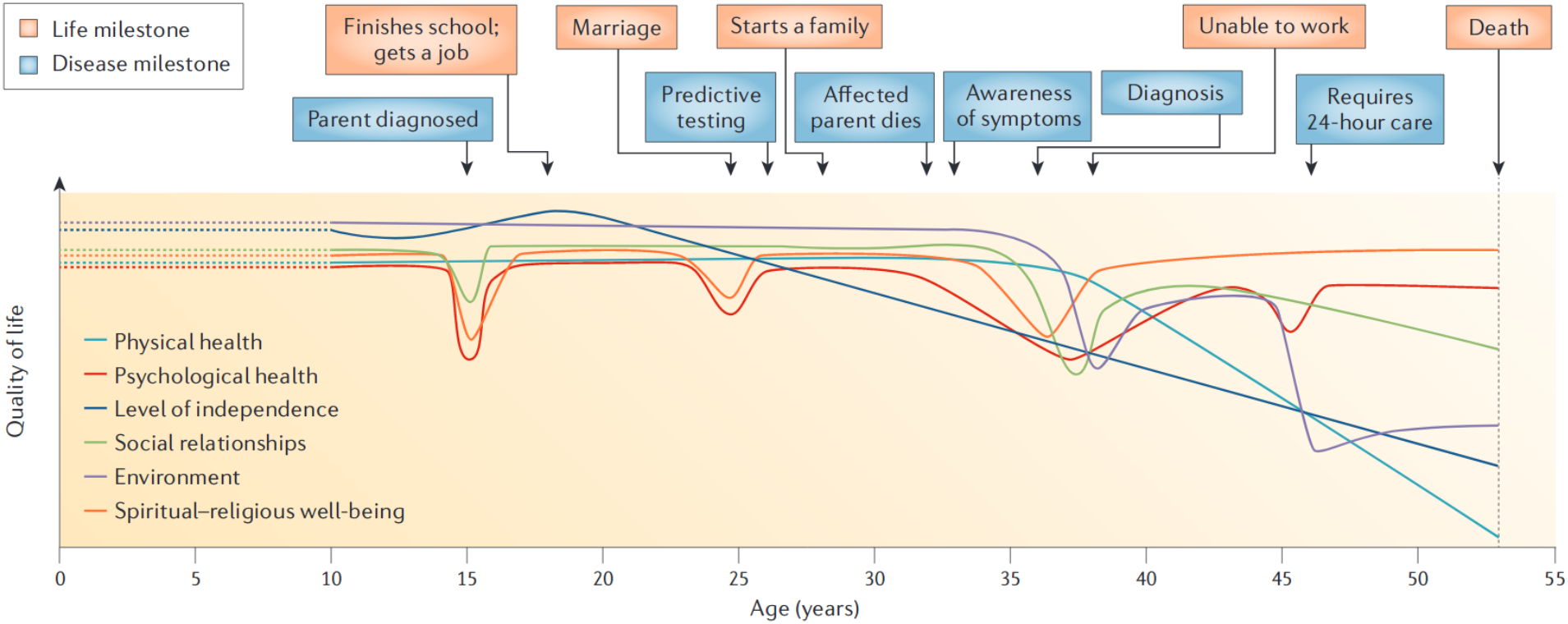


Progressive, hereditary brain disease that causes changes in movement, cognition and behaviour

Autosomal dominant – 50% of inheriting
Fully penetrant – everyone with gene will develop HD

Huntington's disease

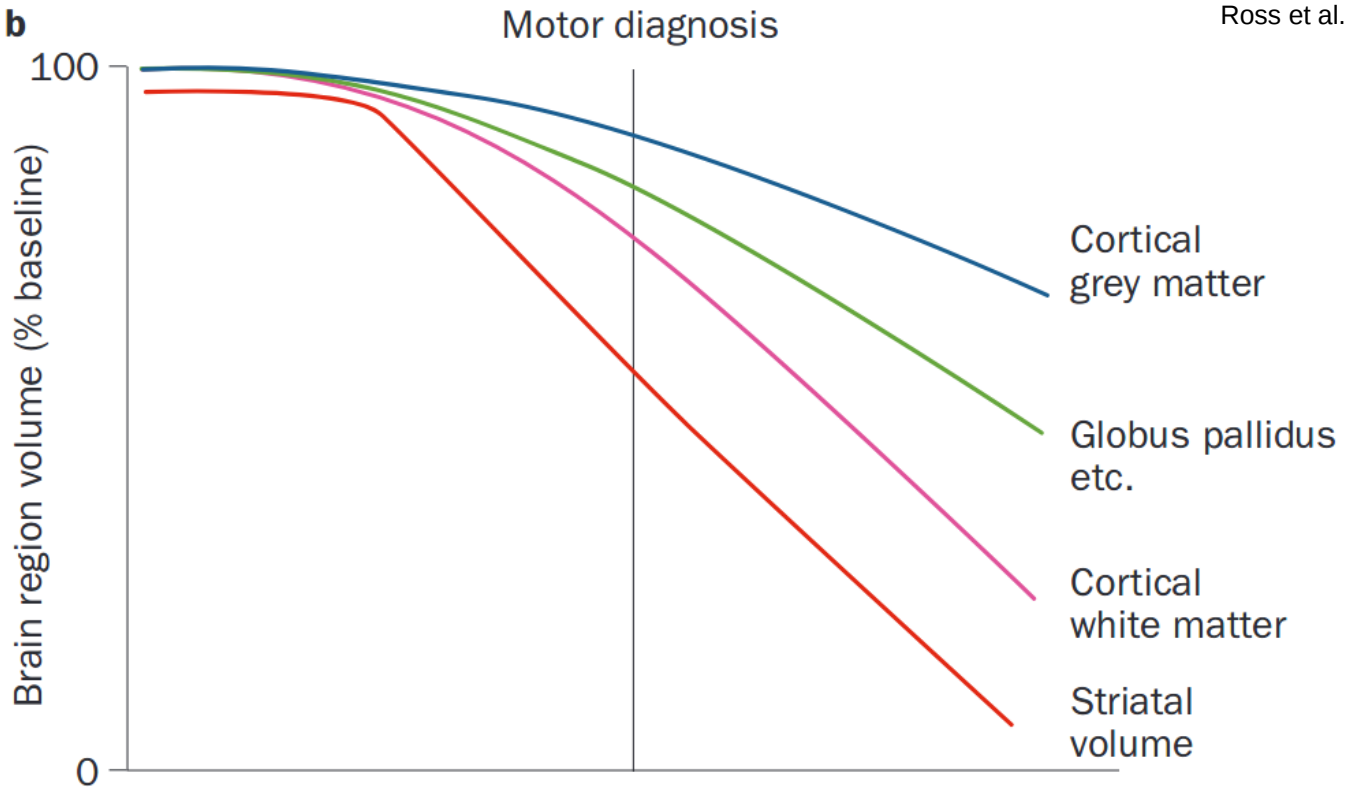
Bates et al. Nature Reviews Disease Primer. 2015



Diagnosis made at onset of movement disorder, typically with chorea and impaired voluntary movement

Huntington's disease

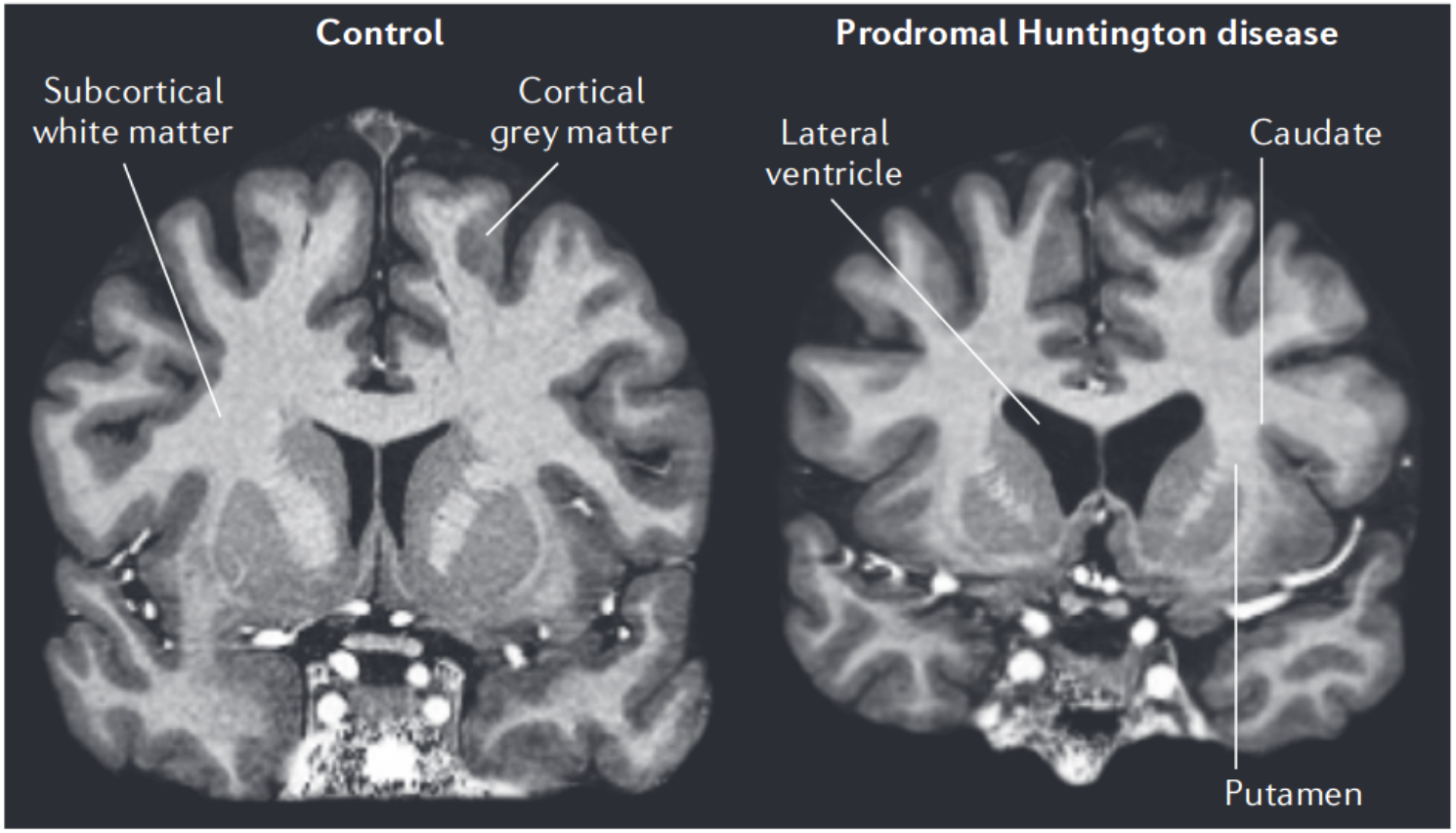
Ross et al. Nature Reviews 2014



Brain changes in HD – specific regions of the brain are atrophied

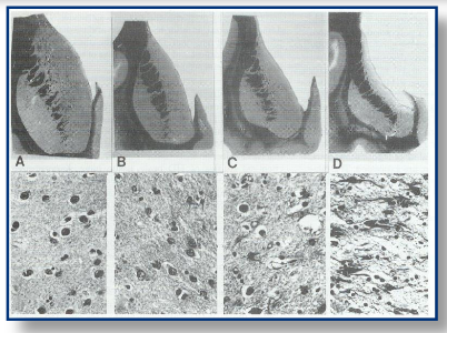
Huntington's disease

Bates et al. Nature Reviews Disease Primer. 2015

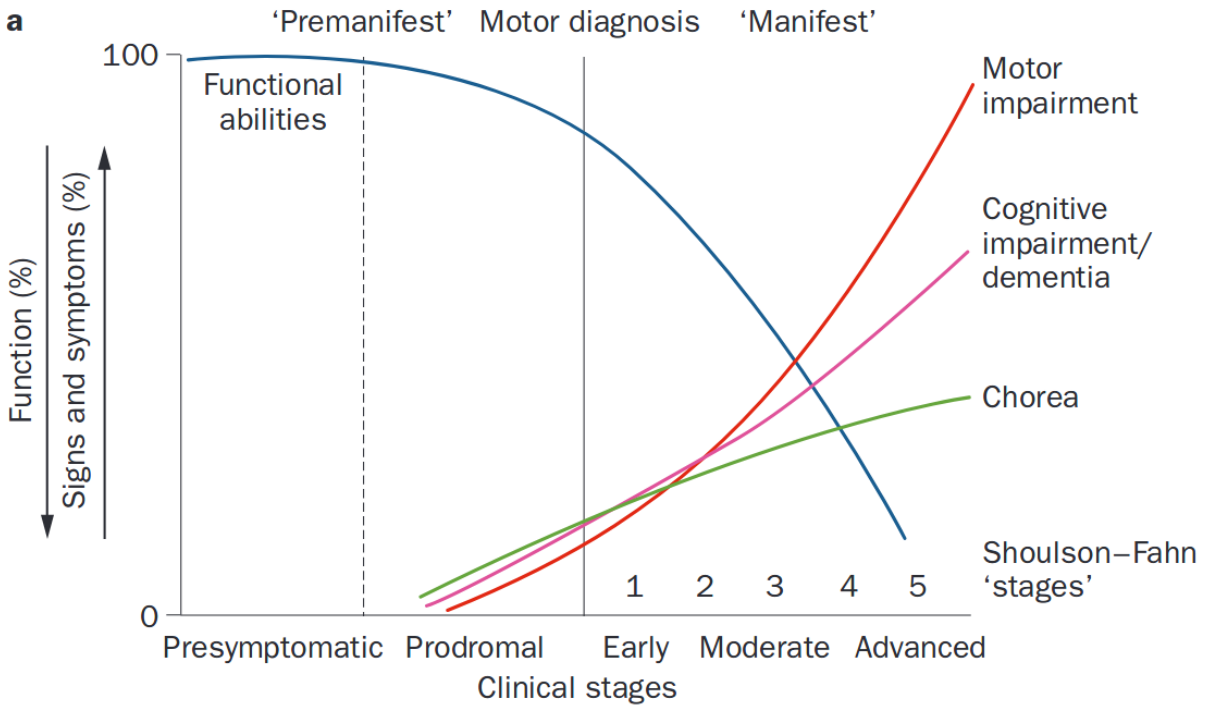


MRI provides spatial intensity measurements that depend on tissue properties

Observed changes reflected by microscopy (histology)



Can we estimate where a patient is along their disease path?

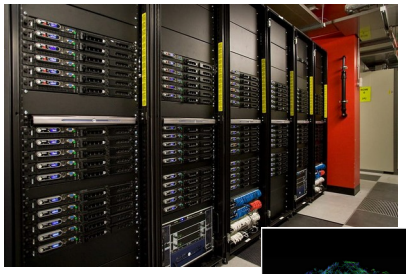


Patient stage is a latent variable – it generates the observed measurements, but is not measured directly (unlike in physics events, where we know time)

→ Infer using statistical and machine learning methods

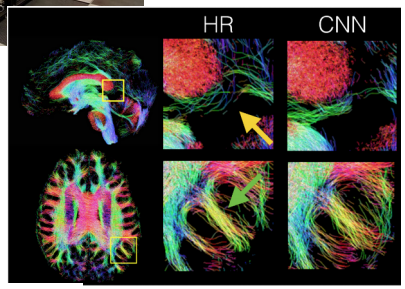
Basic sciences

cmic
Centre for Medical Image Computing



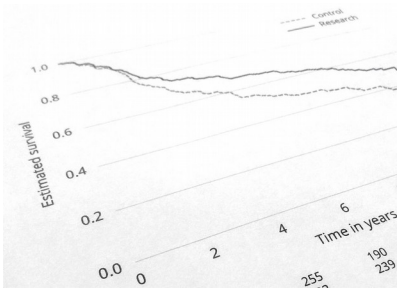
Cluster computing

Imaging + machine learning



UCL EPSRC CDT in Medical Imaging

Statistical methods



pond
Progression Of Neurodegenerative Disease

Clinical sciences

HUNTINGTON'S DISEASE CENTRE

Leonard Wolfson Experimental Neurology Centre

Eisai

Advanced imaging

wellcome centre human neuroimaging

Clinical trials

Some quick definitions

Biomarker: any biological measurement that tracks disease progression

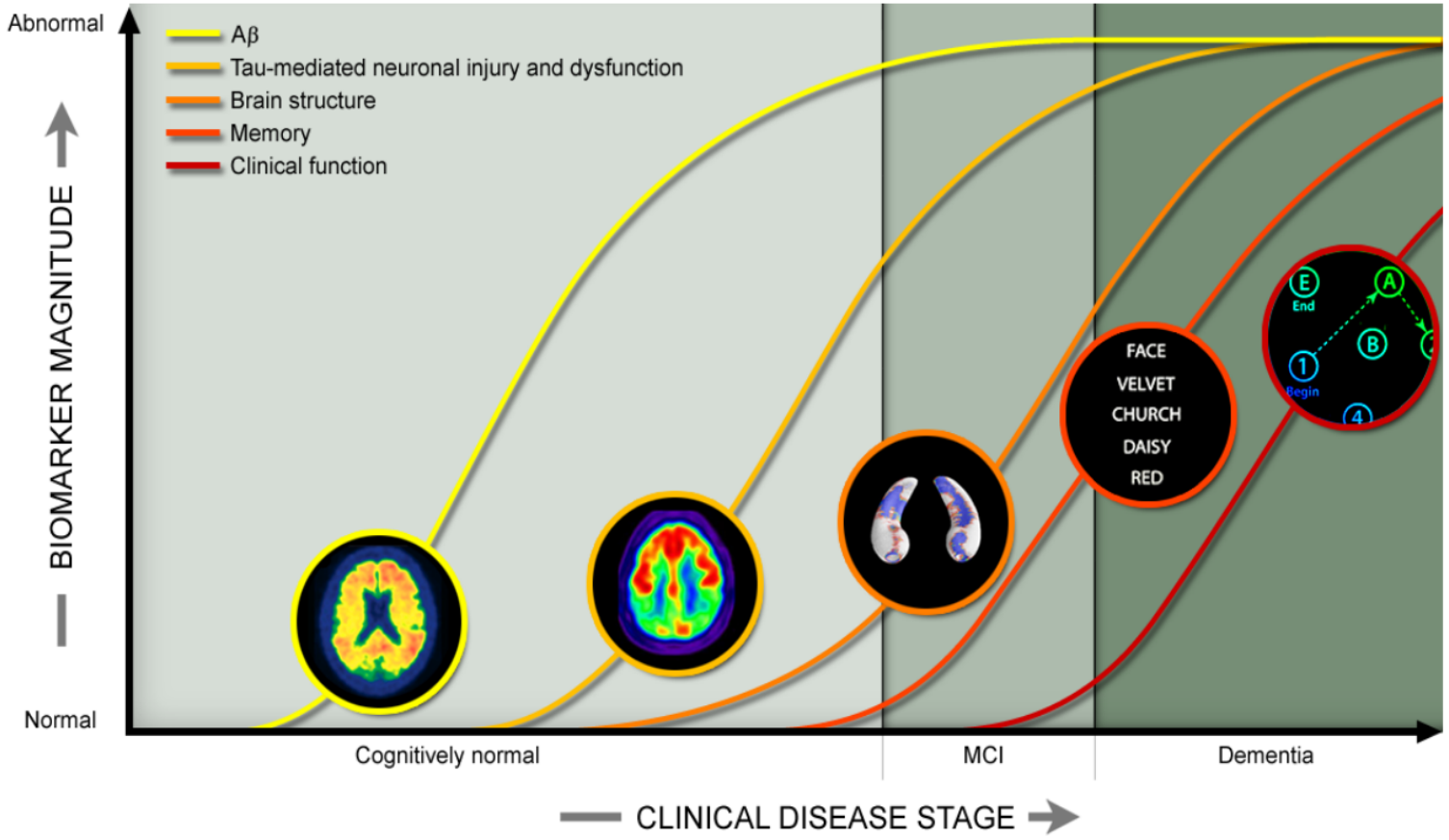
Event: transition of a biomarker from a normal to abnormal state (Markovian)

Sequence: order of events over sample of interest

Cross-sectional: data from a single time-point

High level: Disease progression modelling

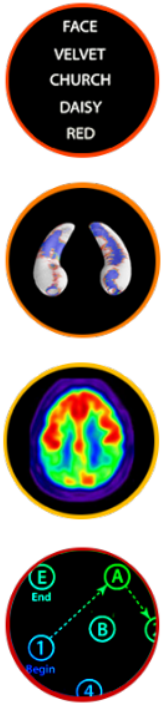
http://adni.loni.usc.edu/study-design/#background-container



A picture of how components of a disease progresses over time

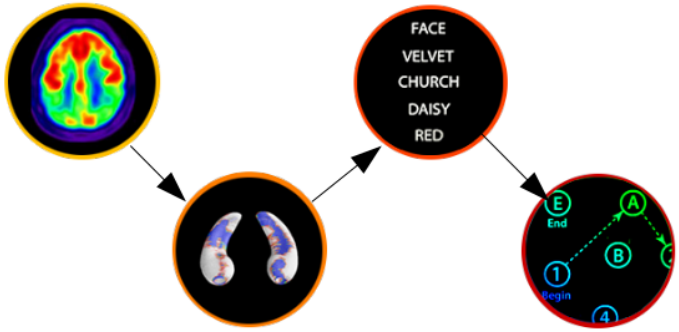
High level: Disease progression modelling

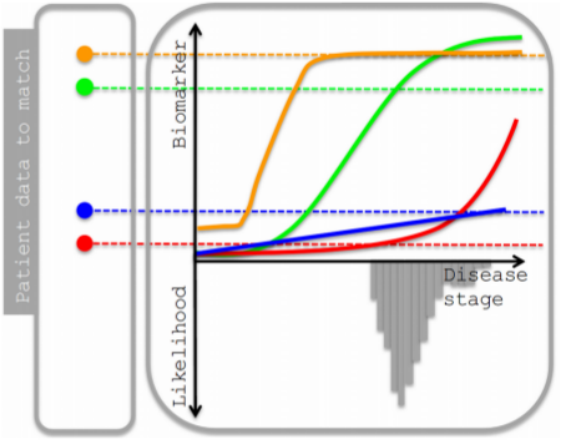
Disease progression models learn patterns of disease-related changes from data



Patient data

Machine learning





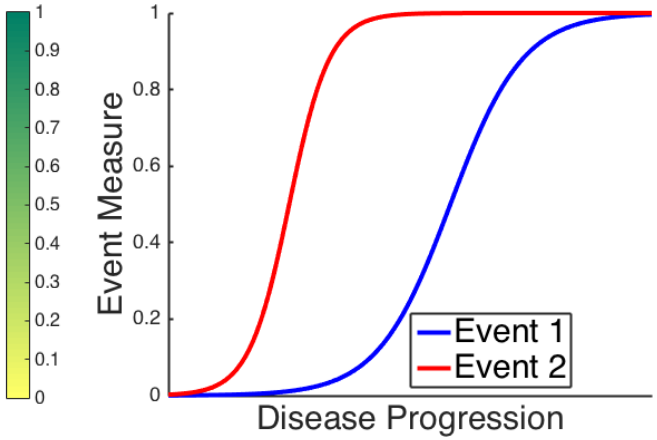
Disease progression model

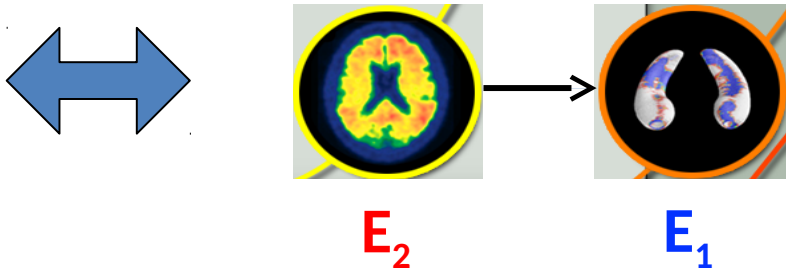
- Can use models to infer temporal ordering of changes
- Can also stage and stratify patients →clinical trial design

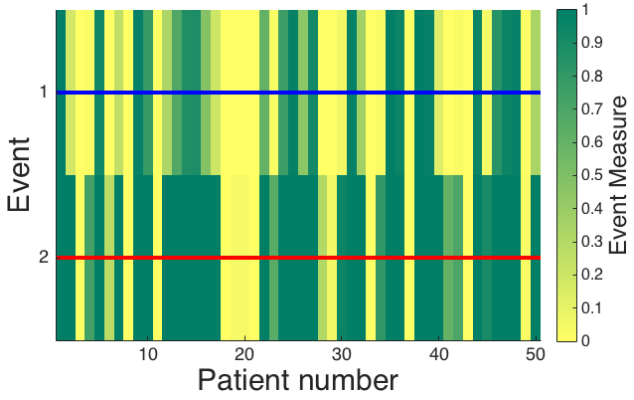
Background: Event-based model (EBM)

EBM estimates ordering of **binary events** from data – normal or abnormal

Data can be cross-sectional and any combination of types (imaging, clinical, genetic...)







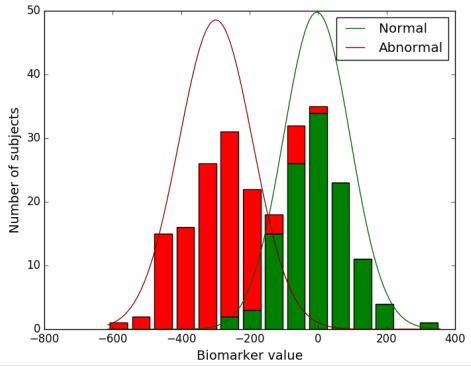
<u>Simple example: 2 event measures</u>

More patients have greater abnormality in Event 2 than Event 1

→Event 2 **measurably abnormal** before Event 1

Background: Event-based model (EBM)

More formally: EBM is a generative model of observed data from unknown sequence



data

uniform prior

$$P(X|S)=\prod_{j=1}^{N}\left[\sum_{k=0}^{Z}\left(P(k)\prod_{i=1}^{k}P(x_{ij}|E_i)\prod_{i=k+1}^{Z}P(x_{ij}|\neg E_i)\right)\right]$$

sequence

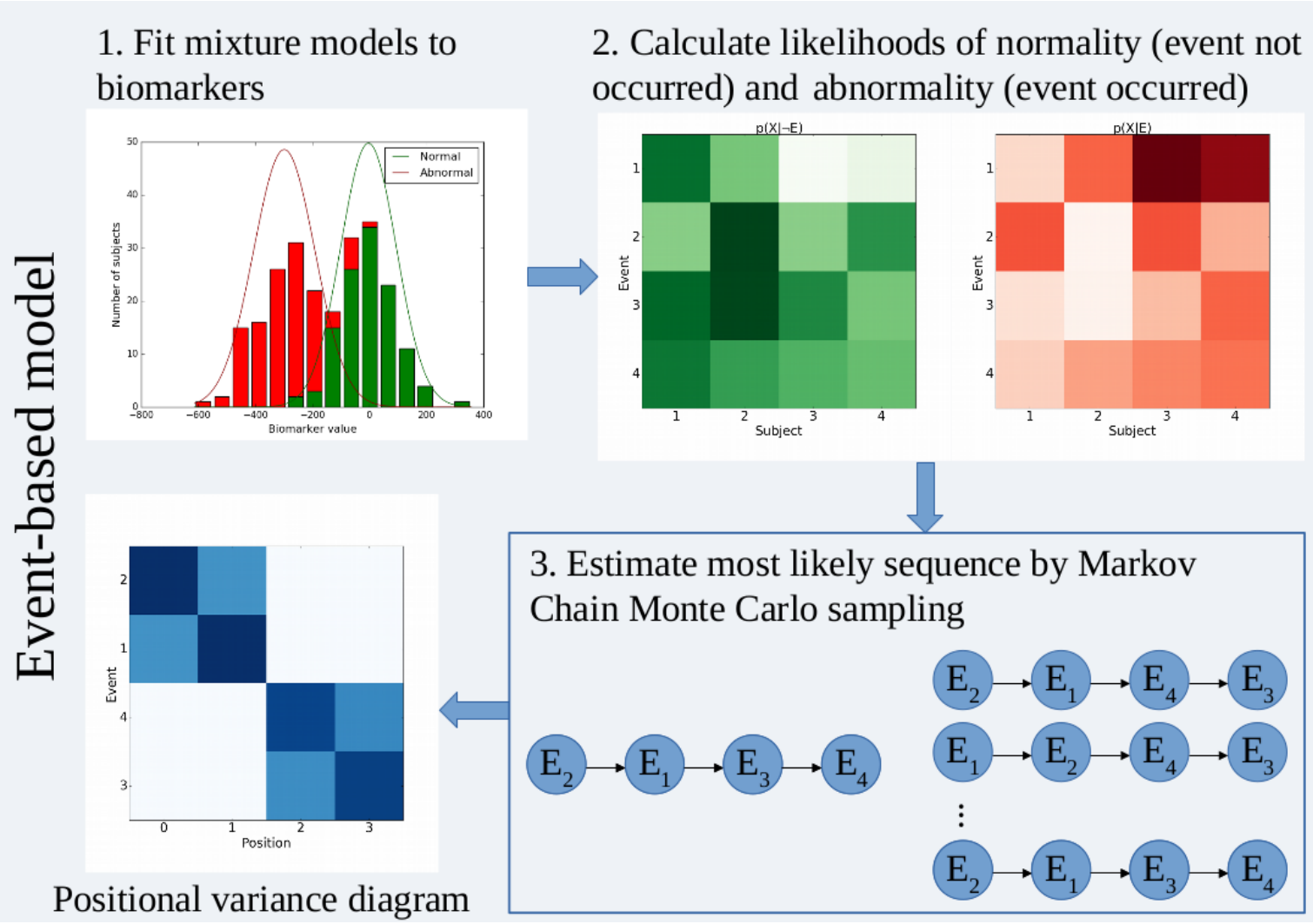
prob. Abnormal

prob. Normal

- The EBM needs likelihood distributions for normal and abnormal subjects

→Learn directly from data

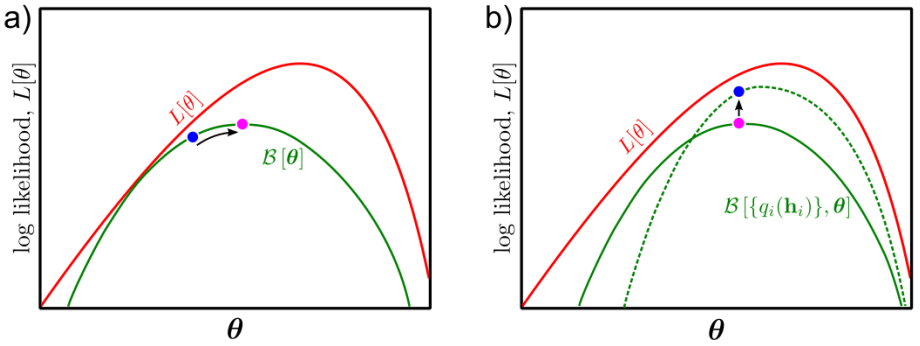
Example: Event-based model (EBM)



Toolkit: parameter estimation

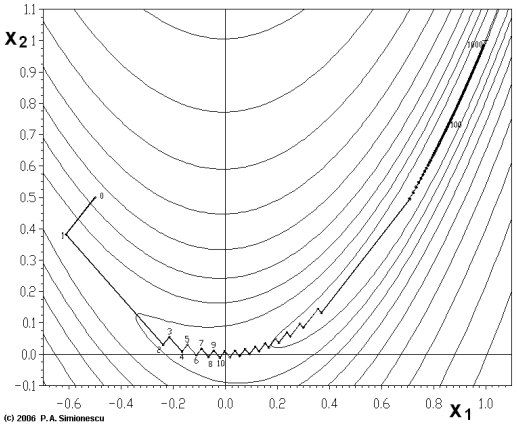
1. Mixture model fitting
– Expectation Maximisation

Prince, SJD. Cambridge University Press. 2012

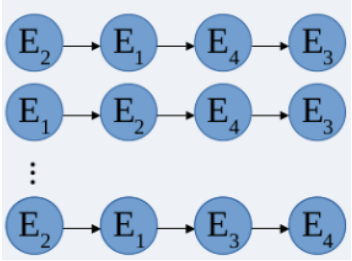


2. Latent variable (sequence) fitting
– Gradient Ascent

wikipedia.org/wiki/gradient_descent



3. Uncertainty estimation
– Markov Chain Monte Carlo



$$a = p(X \mid S')/p(X \mid S_t)$$

Mapping the Huntington's timeline

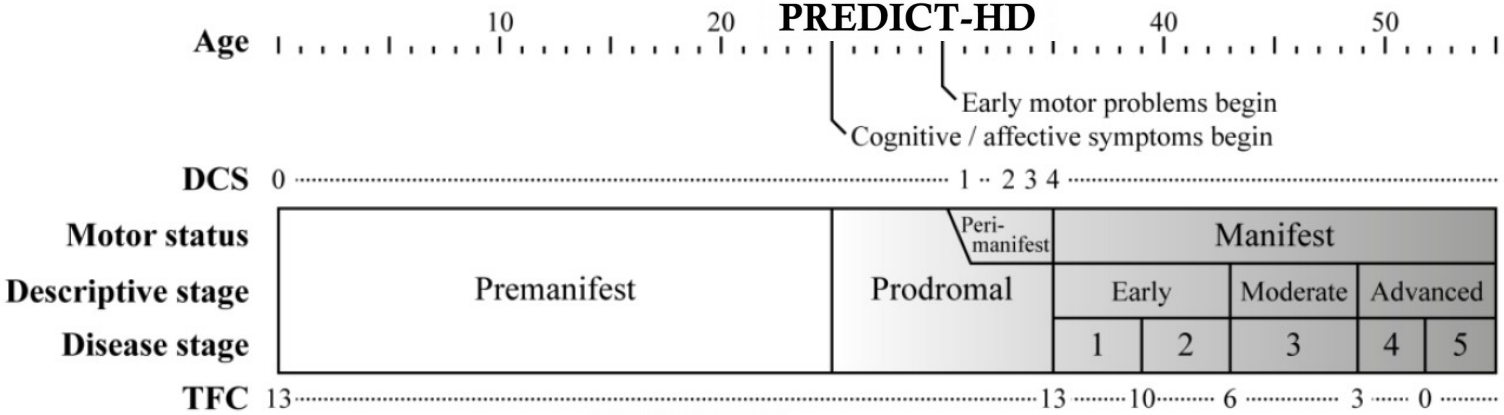
Unique access to the largest combined multi-modal dataset in Huntington's disease

IMAGE-HD

TRACK-HD

TrackOn-HD

PREDICT-HD



Enroll-HD

HD YAS

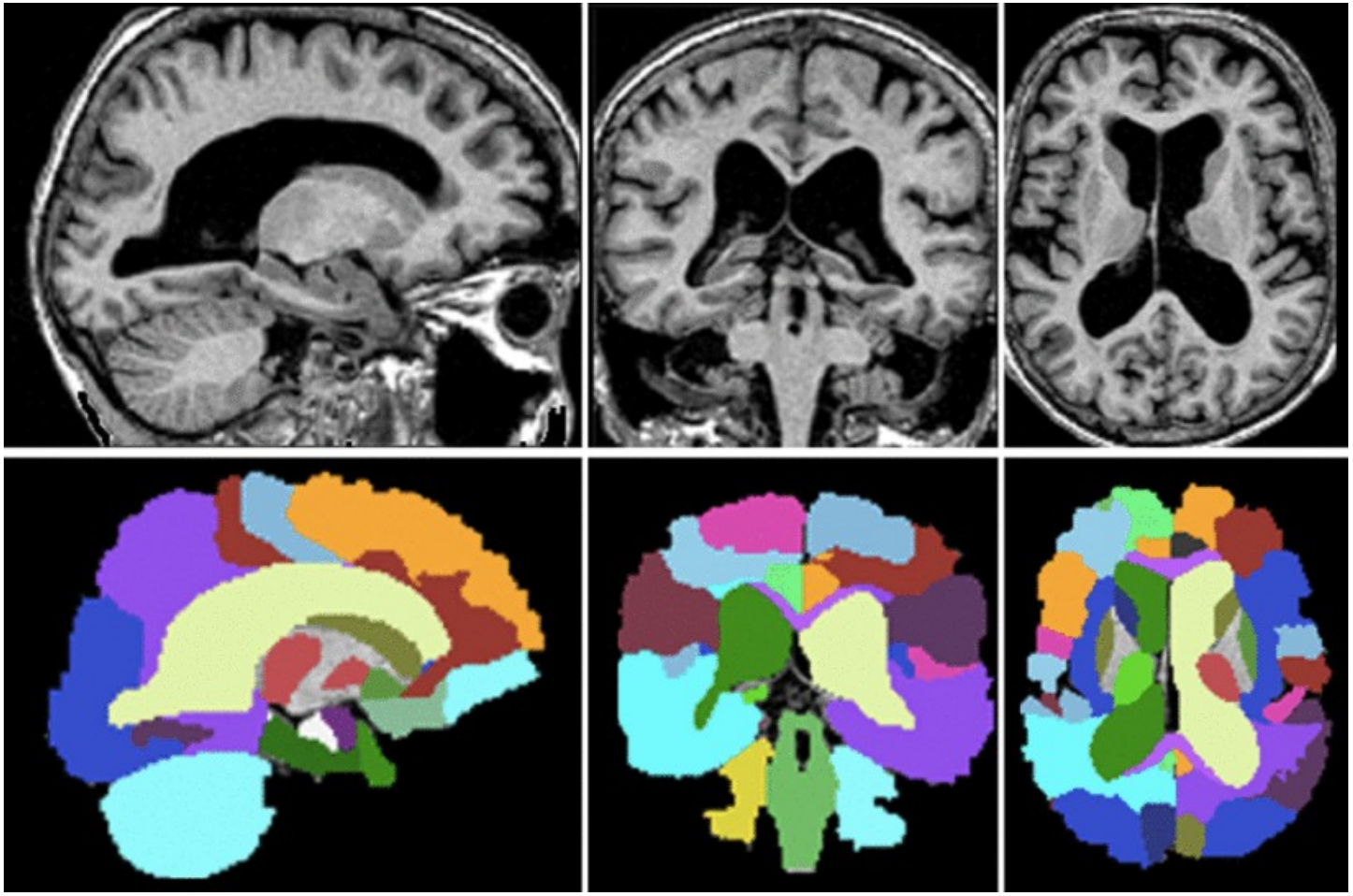
HD CSF

Train and test disease progression models

Extend disease progression models

Evaluate clinical trials model

Methods: Imaging data

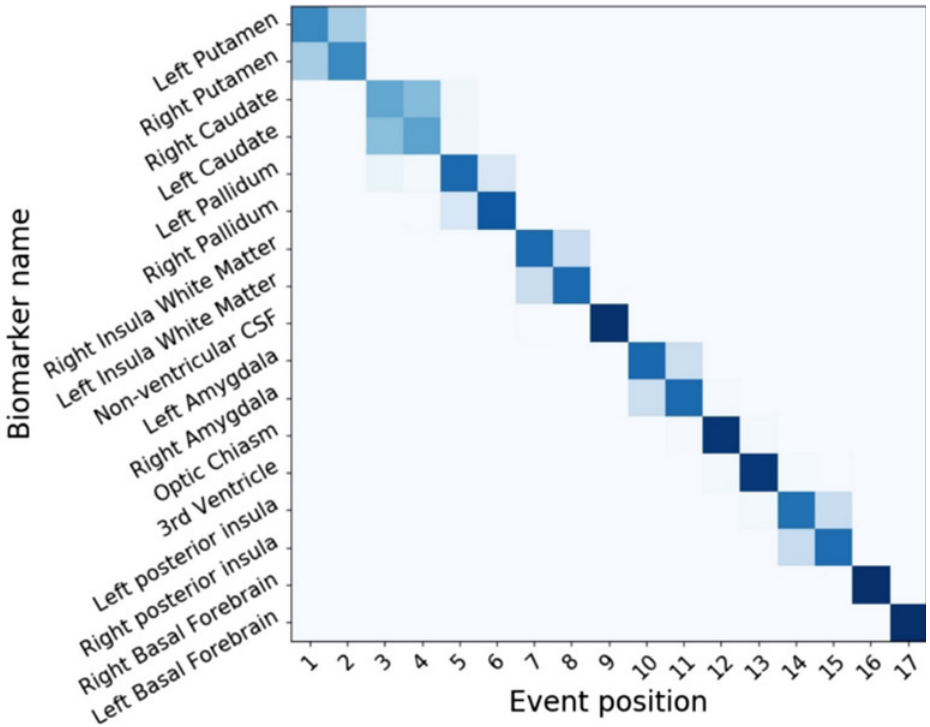


Extract regional brain volumes using Geodesic Information Flows*

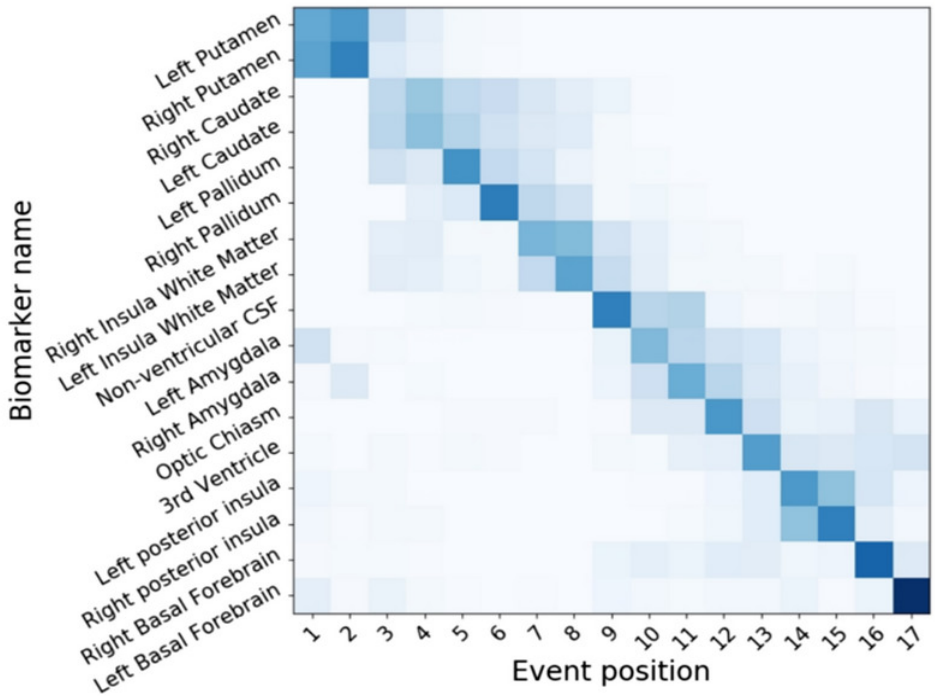
→ Reduces inter-subject variability by using spatially variant graphs to connect morphologically similar subjects

* MJ Cardoso *et al*. Geodesic Information Flows: Spatially-Variant Graphs and Their Application to Segmentation and Fusion. IEEE Transactions on Medical Imaging, 34 (2015), pp. 1976-1988, doi: 10.1109/TMI.2015.2418298

Direct model fit

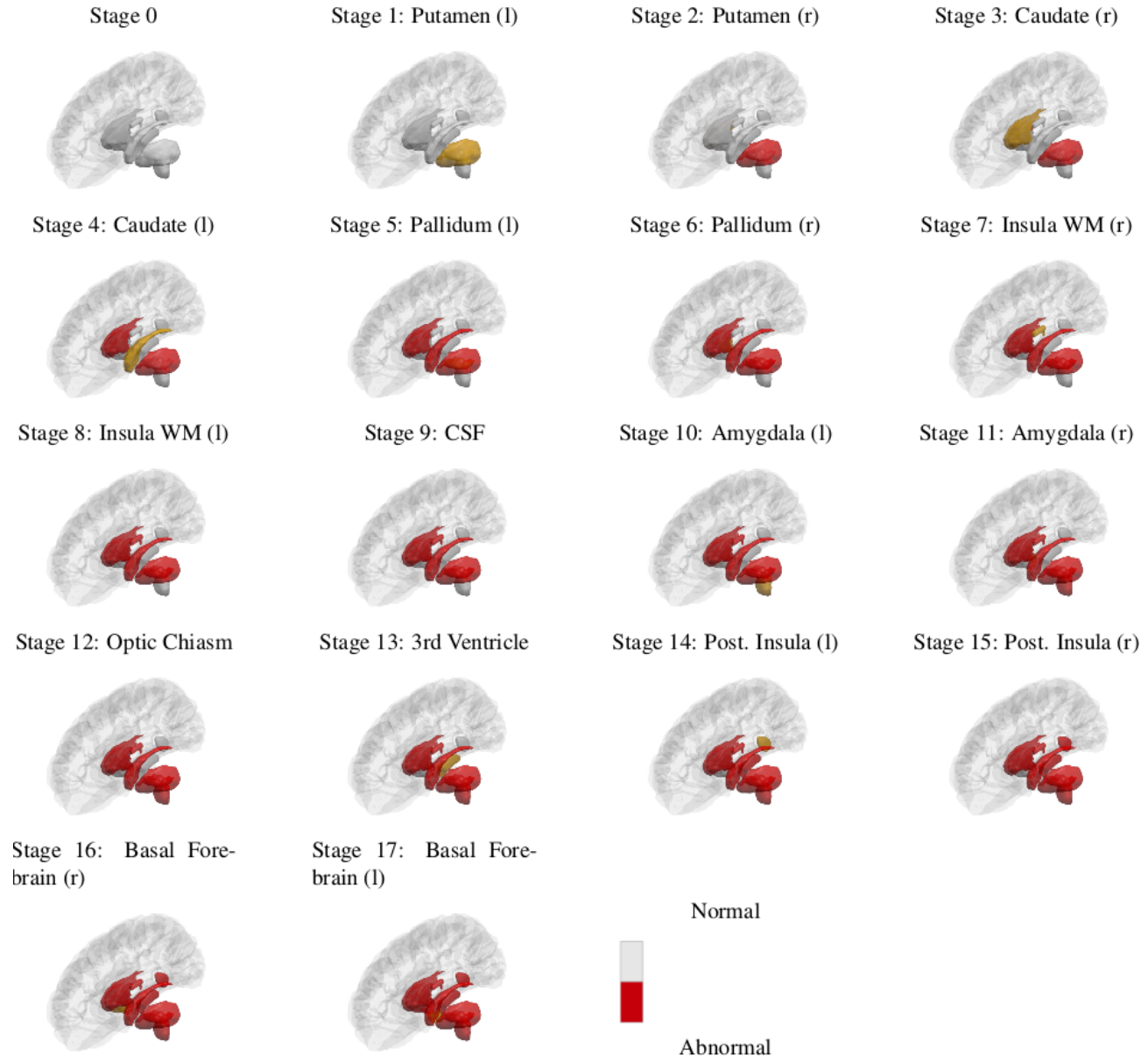


Bootstrapped model fit



- Dark diagonal components indicate strong event ordering
- Lighter indicate possible event permutations

Atrophy progression



Central

HD progression

Peripheral

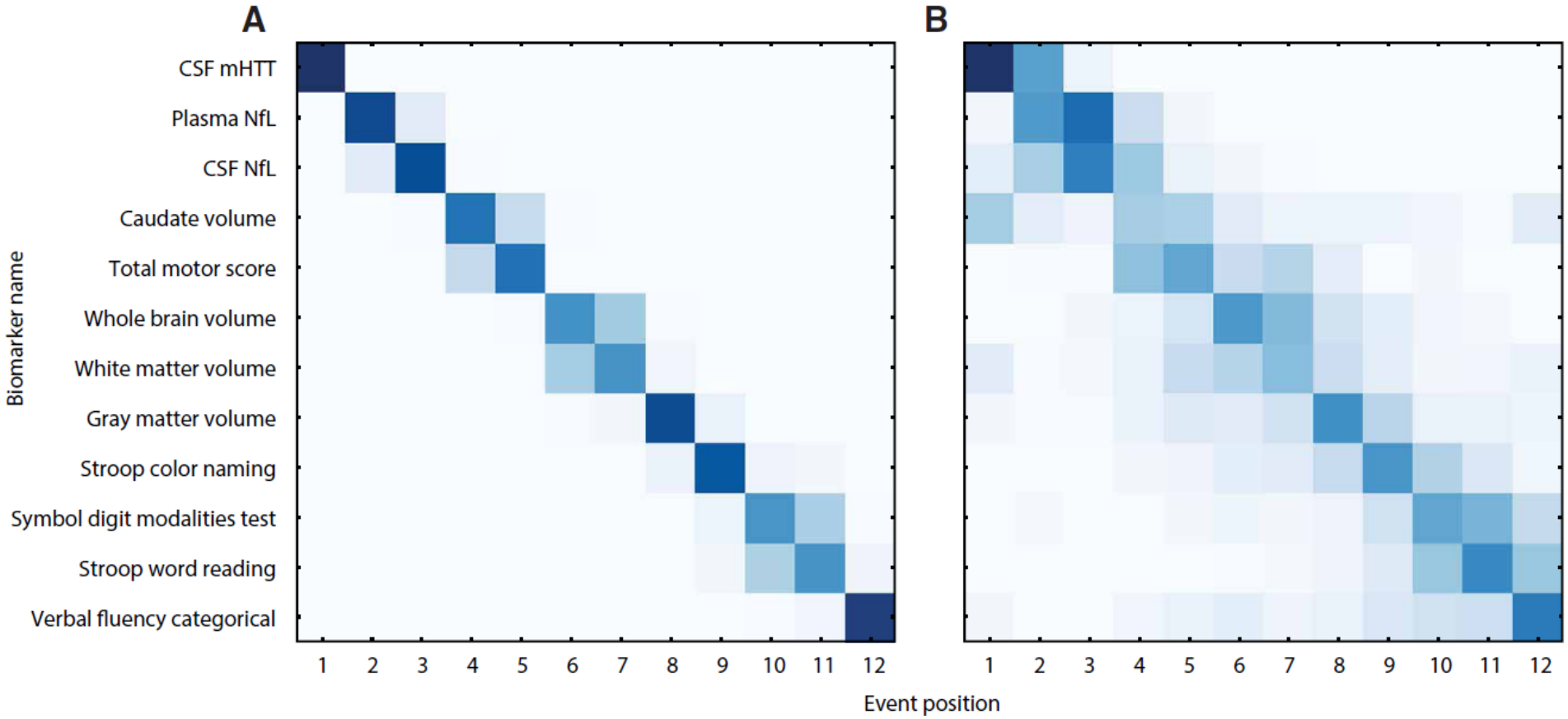
Backup: HD-CSF EBM

SCIENCE TRANSLATIONAL MEDICINE | RESEARCH ARTICLE

HUNTINGTON'S DISEASE

Evaluation of mutant huntingtin and neurofilament proteins as potential markers in Huntington's disease

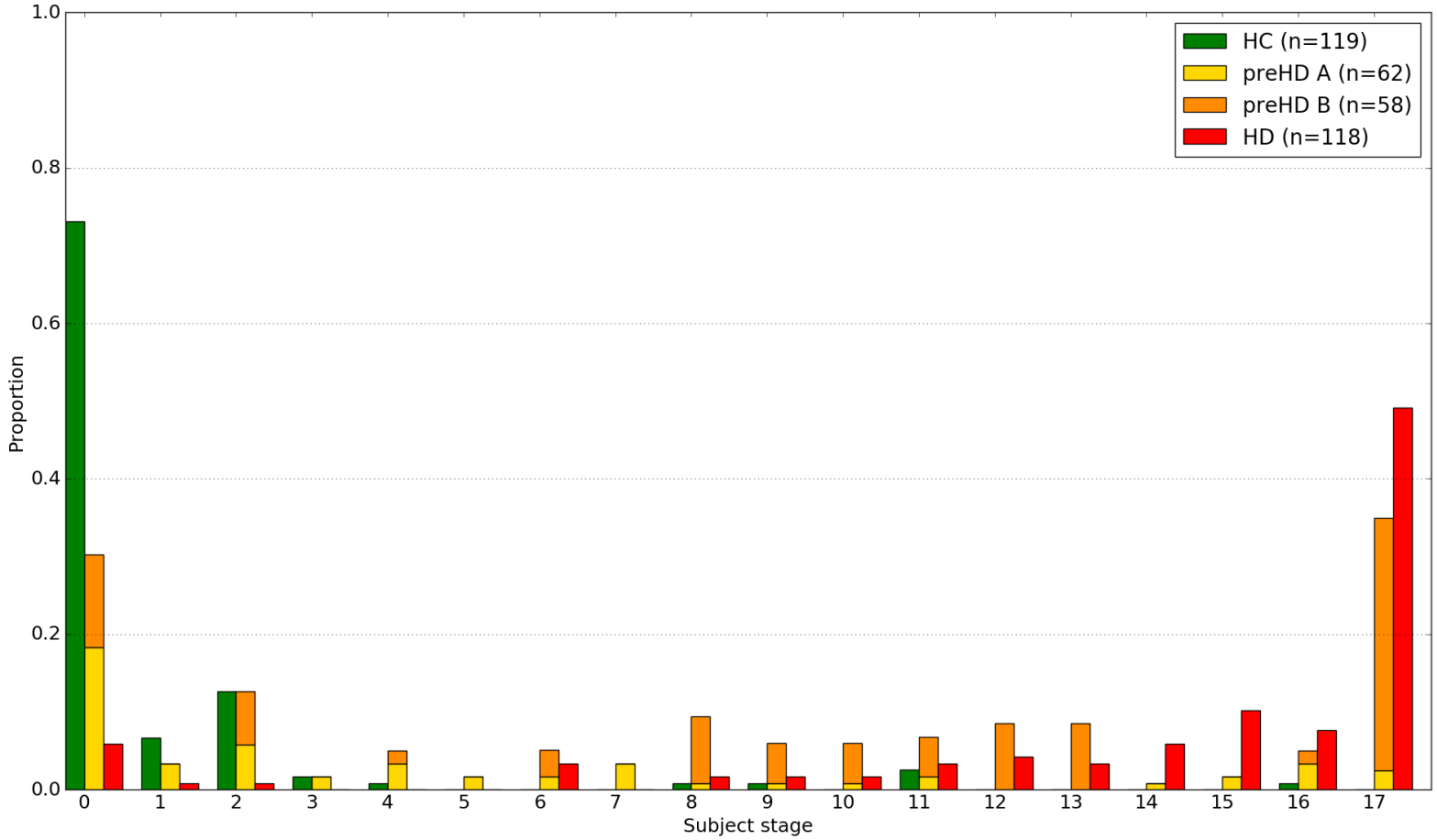
Lauren M. Byrne[1*†], Filipe B. Rodrigues[1†], Eileanor B. Johnson[1], Peter A. Wijeratne[2], Enrico De Vita[3,4], Daniel C. Alexander[2,5], Giuseppe Palermo[6], Christian Czech[6], Scott Schobel[6], Rachael I. Scahill[1], Amanda Heslegrave[7], Henrik Zetterberg[7,8,9,10], Edward J. Wild[1*]



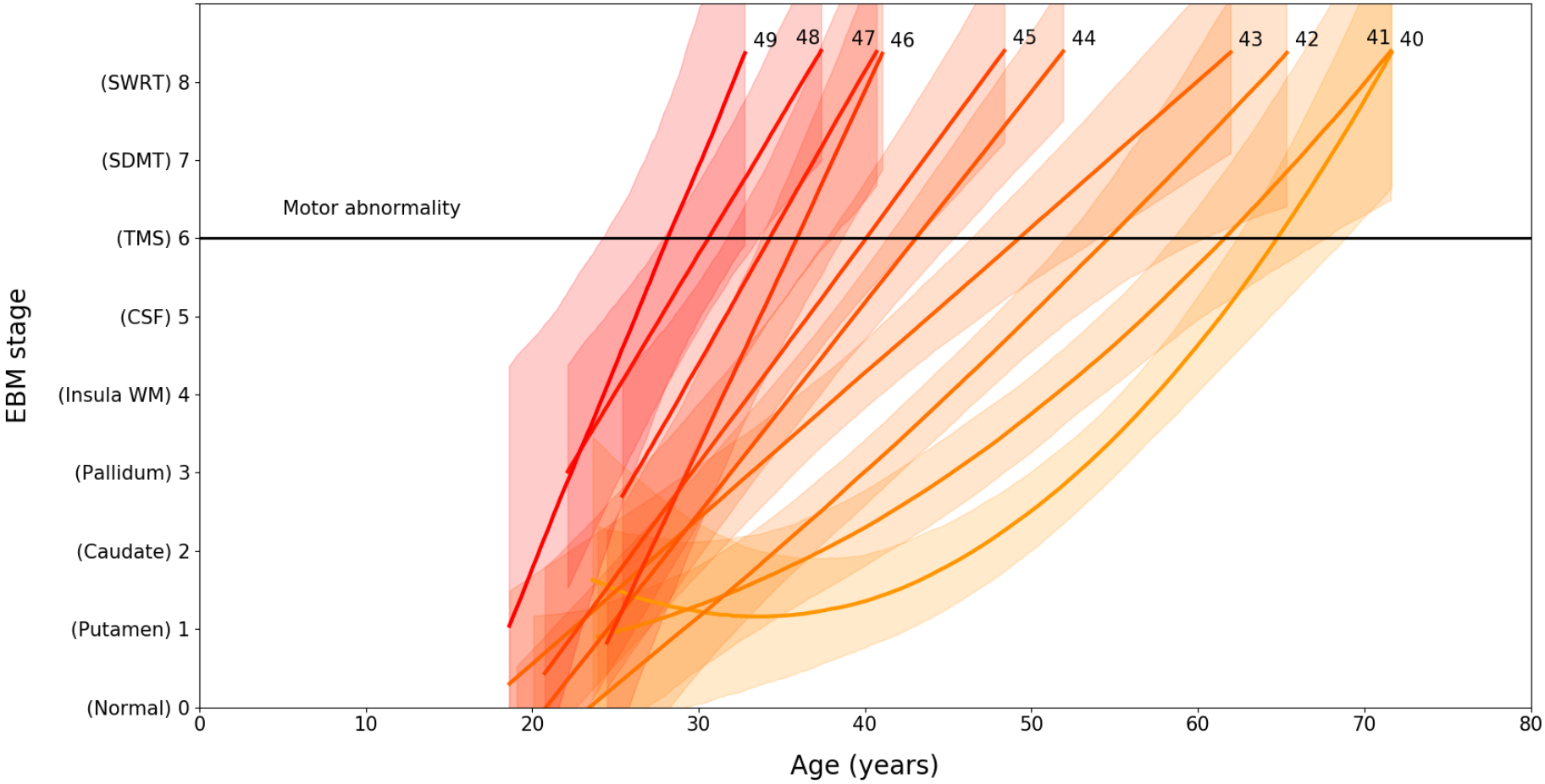
- Biofluid markers change before imaging and clinical markers

Simplest way is to take the stage that maximises the likelihood for each patient

$$argmax_k P(X_j|\bar{S},k) = argmax_k P(k) \prod_{i=1}^{k} P(x_{ij}|E_i) \prod_{i=k+1}^{I} P(x_{ij}|\neg E_i)$$

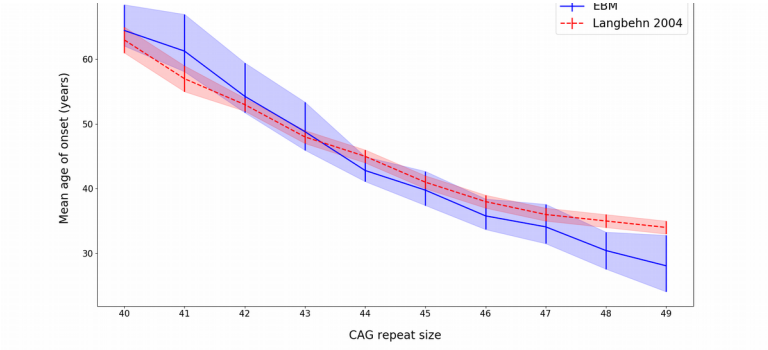


Extending EBM-HD + cross-validation



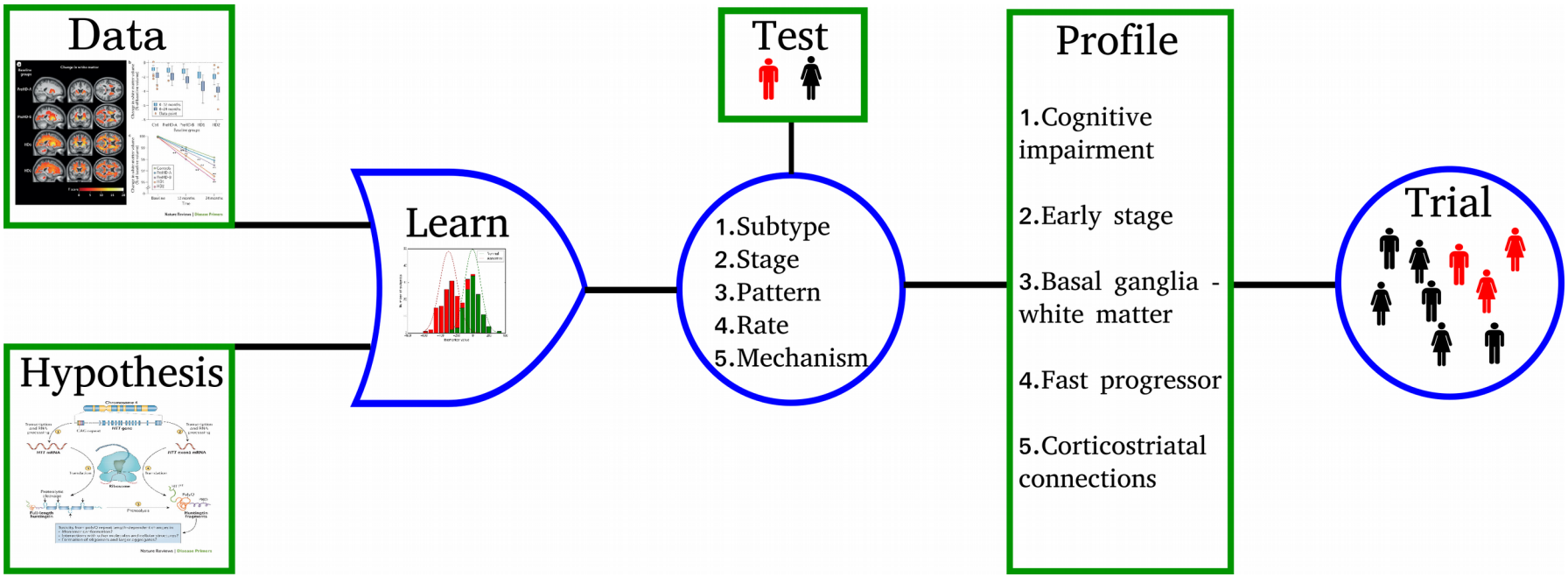
- Estimate age at event e.g.
  for CAG 40, WM atrophy at ~60 years old
  for CAG 49, WM atrophy at ~25 years old

- Age of onset agrees well with gold standard



Big vision: computational models for clinical trials

Patient data + machine learning = personalised profiles for clinical trial design



Model can be used for both prospective and retrospective analysis

→ Save money and time
→ Optimise trial design

We've looked at cross-sectional modelling, so the natural next step is...

Longitudinal modelling of Huntington's disease biomarkers

- Use Gaussian Process Progression model with Huntington's disease data
  - Biomarker trajectories, relative ordering
- Explore potential methodological developments
  - 'Subtyping' (i.e. clustering covariance)
- Other potential applications of GP-based models in HD
  - GPs with mechanistic constraints, voxelwise data, VAEs
  - Suggestions very welcome!