

Data-driven Disease Progression Modelling: thinking outside the black box

Neil P. Oxtoby

Principal Research Fellow (research track Assoc. Prof.)
UCL Centre for Medical Image Computing



Data Science Lead, DEMON Network
demon dementia.com



Today (Tonight)

- Background & Motivation: Alzheimer's disease (+ others)
 - Lack of a well-defined and consistent “disease time” axis
- Data-driven Disease Progression Modelling
 - Math + Human Insight + ML + “Big” Data

This talk is based on two papers

2017



Imaging plus X: multimodal models of neurodegenerative disease

Neil P. Oxtoby and Daniel C. Alexander, for the EuroPOND consortium

nature reviews neuroscience

2024

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young ^{1,2,6} , Neil P. Oxtoby ^{1,6} , Sara Garbarino ³, Nick C. Fox⁴, Frederik Barkhof ^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander ¹

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young ^{1,2,6} , Neil P. Oxtoby ^{1,6} , Sara Garbarino ³, Nick C. Fox⁴, Frederik Barkhof ^{1,5},
Jonathan M. Schott⁴ & Daniel C. Alexander ¹

Definition of “Data-driven disease progression model”:

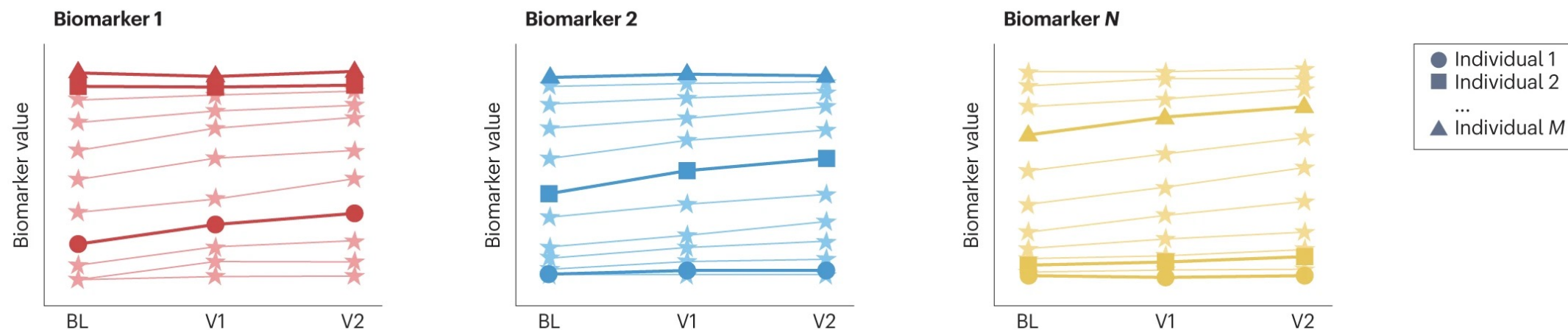
- Constructs a quantitative timeline of disease
- Directly informed by measured data

- Construct a quantitative timeline of disease.
- Directly informed by measured data

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young^{1,2,6}, Neil P. Oxtoby^{1,6}, Sara Garbarino³, Nick C. Fox⁴, Frederik Barkhof^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander¹

Disease biomarker data indexed by visit

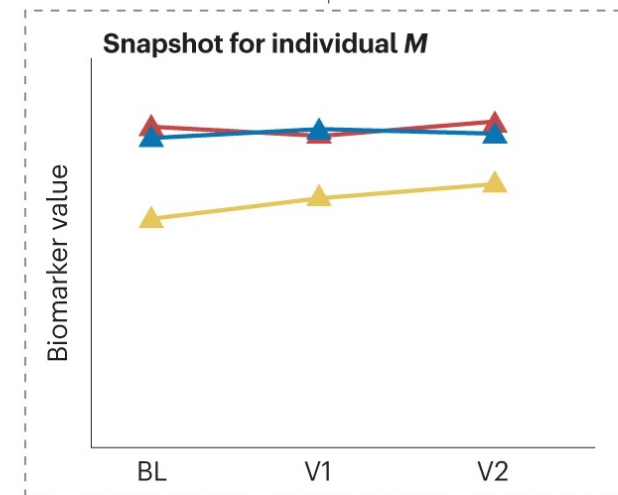
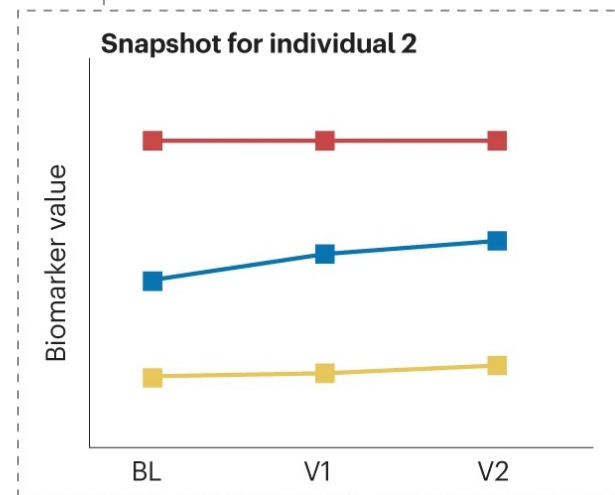
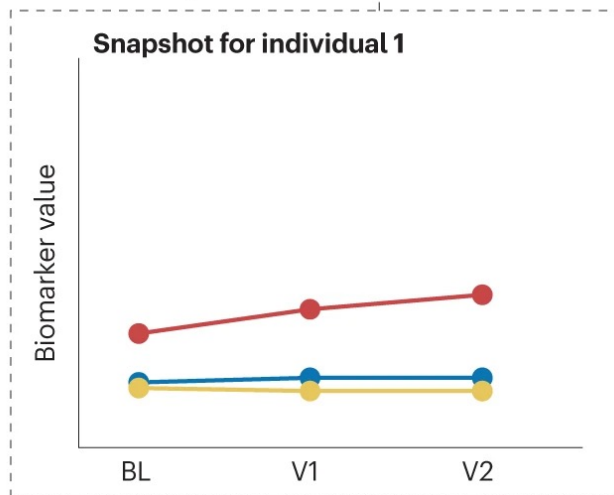


- Construct a quantitative timeline of disease.
- Directly informed by measured data

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young^{1,2,6}, Neil P. Oxtoby^{1,6}, Sara Garbarino³, Nick C. Fox⁴, Frederik Barkhof^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander¹

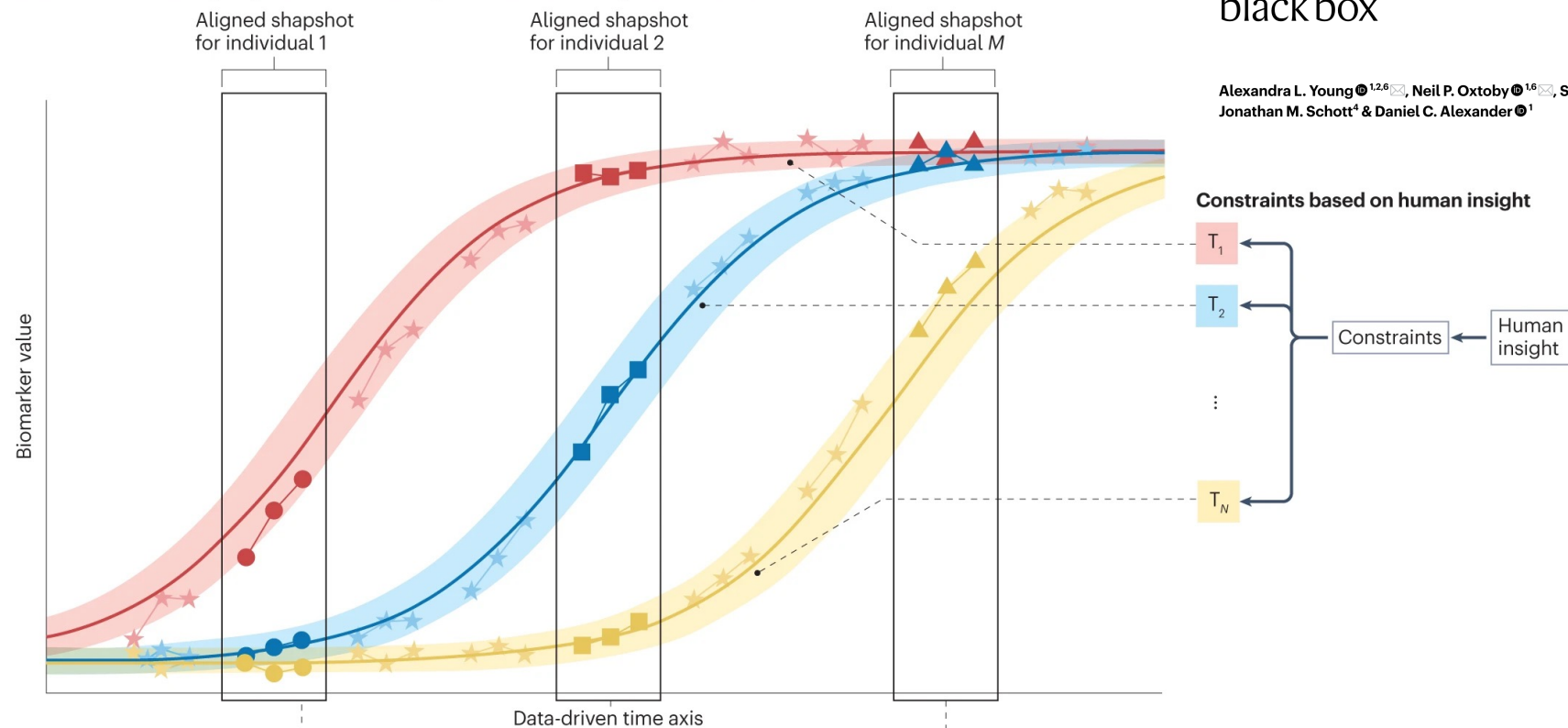
Short-term snapshots



- Construct a quantitative timeline of disease.
- Directly informed by measured data

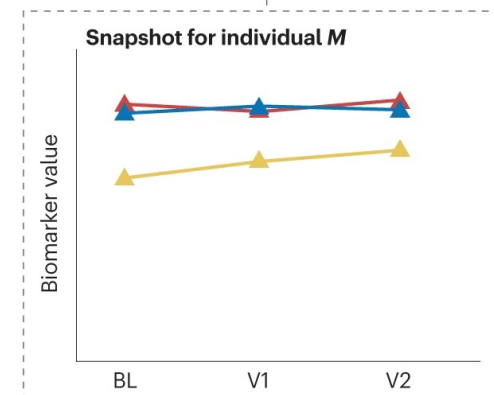
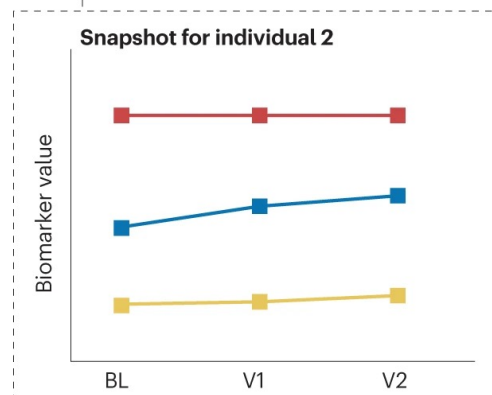
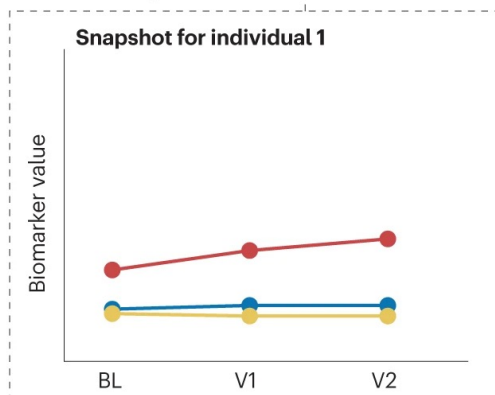
Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Data-driven disease progression modelling reconstructs long-term disease timelines



Alexandra L. Young^{1,2,6}, Neil P. Oxtoby^{1,6}, Sara Garbarino³, Nick C. Fox⁴, Frederik Barkhof^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander¹

Short-term snapshots



Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young ^{1,2,6} , Neil P. Oxtoby ^{1,6} , Sara Garbarino ³, Nick C. Fox⁴, Frederik Barkhof ^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander ¹

Taxonomy:

- Phenomenological
- Pathophysiological, a.k.a., Mechanistic

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young ^{1,2,6} , Neil P. Oxtoby ^{1,6} , Sara Garbarino ³, Nick C. Fox⁴, Frederik Barkhof ^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander ¹

Taxonomy:

- “Top-down”: Phenomenological
- “Bottom-Up”: Pathophysiological, a.k.a., Mechanistic

What do we know about Alzheimer's?

- Defined by *post mortem* histopathology
 - Braak staging
- Clinical syndrome: memory *etc.*
- *Looooong* pre-symptomatic period: decades of pathology
 - Virtually impossible to identify future patients
 - Risk factors: genetics, etc.
 - Rare familial/inherited forms
- Heterogeneity in syndrome, onset, progression, and pathology!
 - **Imaging** can probe pathology *in vivo* (PET, MRI)



What do we know about Alzheimer's?

- Defined by *post mortem* histopathology
 - Braak staging
- Clinical syndrome: memory *etc.*
- *Looooong* pre-symptomatic period: decades of pathology
 - Virtually impossible to identify future patients
 - Risk factors: genetics, etc.
 - Rare familial/inherited forms
- Heterogeneity in syndrome, onset, progression, and pathology!
 - **Imaging** can probe pathology *in vivo* (PET, MRI)



Treatments for Alzheimer's?

Treatments for Alzheimer's?

- Amyloid cascade hypothesis (Hardy/Higgins 1992; Selkoe/Hardy 2016)
 - + Plenty of supporting evidence
 - Anti-amyloid therapies: sketchy efficacy in large clinical trials

Treatments for Alzheimer's?

- Amyloid cascade hypothesis (Hardy/Higgins 1992; Selkoe/Hardy 2016)
 - + Plenty of supporting evidence
 - Anti-amyloid therapies: sketchy efficacy in large clinical trials
- Why are clinical trials “failing”? (hundreds since 2003)
 - Too late? (*wrong* time: prevention vs cure)

Treatments for Alzheimer's?

- Amyloid cascade hypothesis (Hardy/Higgins 1992; Selkoe/Hardy 2016)
 - + Plenty of supporting evidence
 - Anti-amyloid therapies: sketchy efficacy in large clinical trials
- Why are clinical trials “failing”? (hundreds since 2003)
 - Too late? (*wrong* time: prevention vs cure)
 - Individual variability? (*wrong* people)

Treatments for Alzheimer's?

- Amyloid cascade hypothesis (Hardy/Higgins 1992; Selkoe/Hardy 2016)
 - + Plenty of supporting evidence
 - Anti-amyloid therapies: sketchy efficacy in large clinical trials
- Why are clinical trials “failing”? (hundreds since 2003)
 - Too late? (*wrong* time: prevention vs cure)
 - Individual variability? (*wrong* people)
 - Insufficient duration?

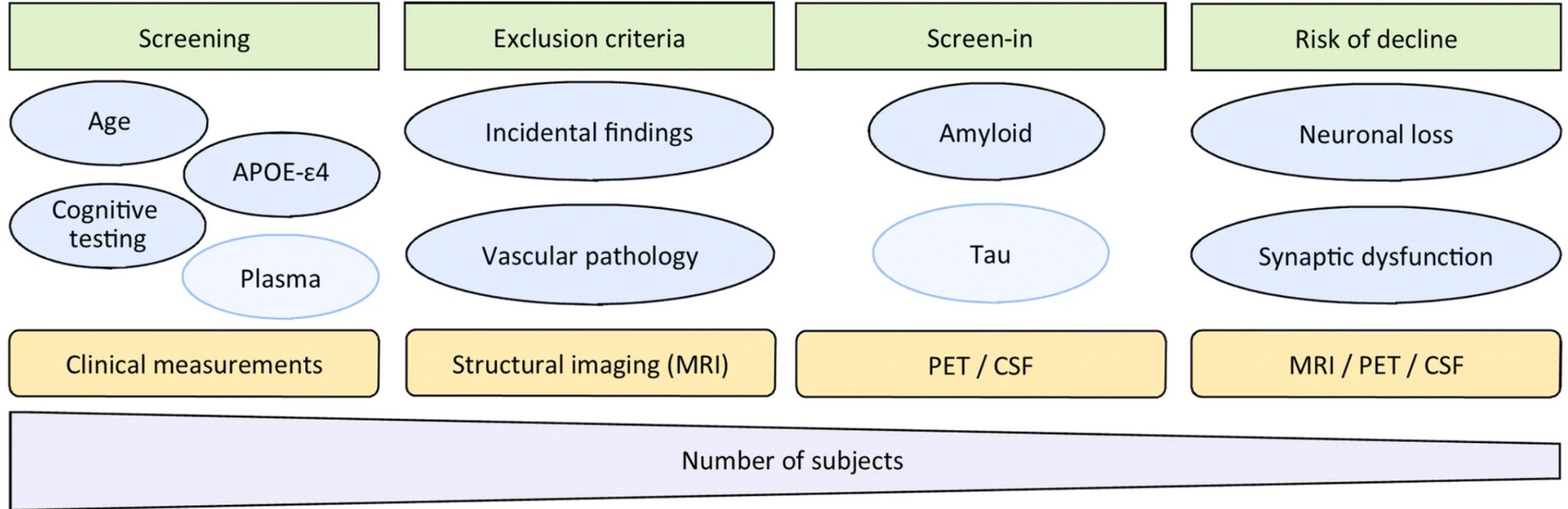
Treatments for Alzheimer's?

- Amyloid cascade hypothesis (Hardy/Higgins 1992; Selkoe/Hardy 2016)
 - + Plenty of supporting evidence
 - Anti-amyloid therapies: sketchy efficacy in large clinical trials
- Why are clinical trials “failing”? (hundreds since 2003)
 - Too late? (*wrong* time: prevention vs cure)
 - Individual variability? (*wrong* people)
 - Insufficient duration?
 - Insensitive end-points? (biology/biomarkers vs clinical benefit)

Treatments for Alzheimer's?

- Amyloid cascade hypothesis (Hardy/Higgins 1992; Selkoe/Hardy 2016)
 - + Plenty of supporting evidence
 - Anti-amyloid therapies: sketchy efficacy in large clinical trials
- Why are clinical trials “failing”? (hundreds since 2003)
 - Too late? (*wrong* time: prevention vs cure)
 - Individual variability? (*wrong* people)
 - Insufficient duration?
 - Insensitive end-points? (biology/biomarkers vs clinical benefit)
 - Wrong target? (*wrong* biology / comorbidities / multitarget strategies)(Salloway, CTAD 2019; Aisen, CTAD 2019)

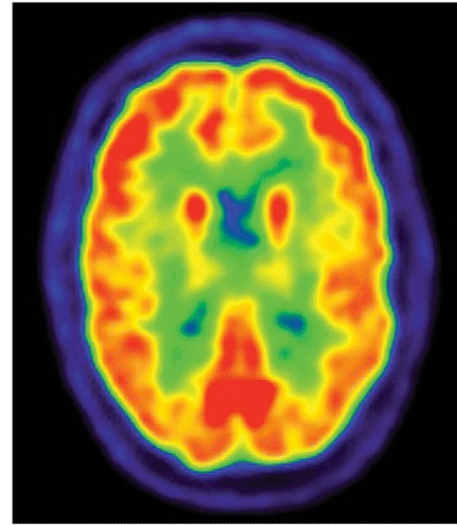
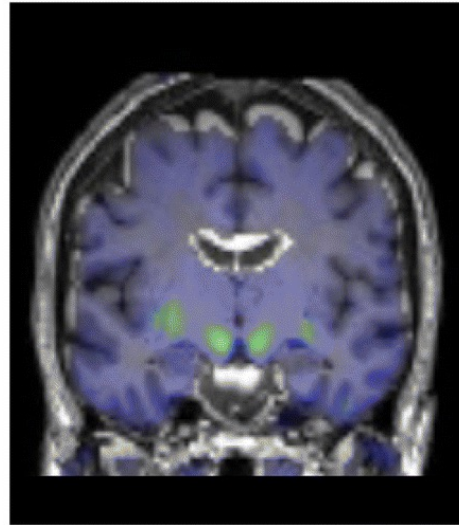
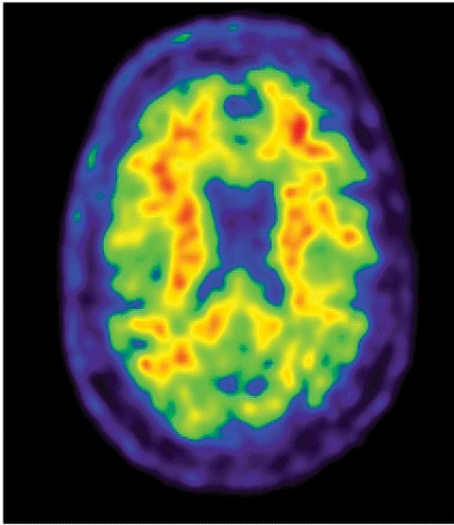
What have clinical trials done?



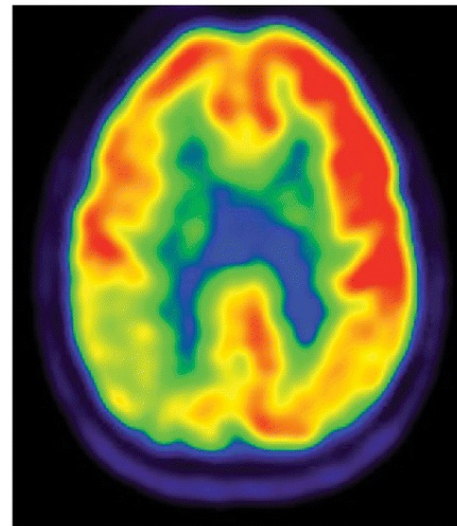
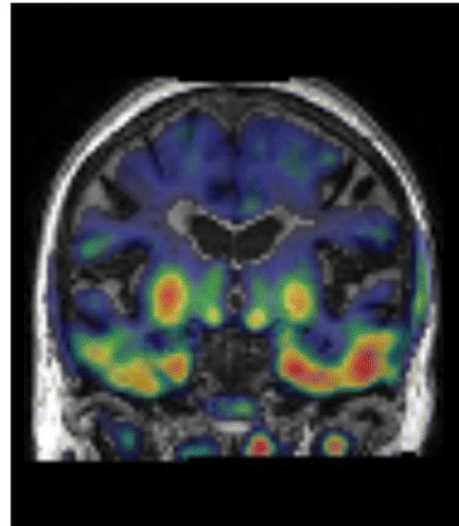
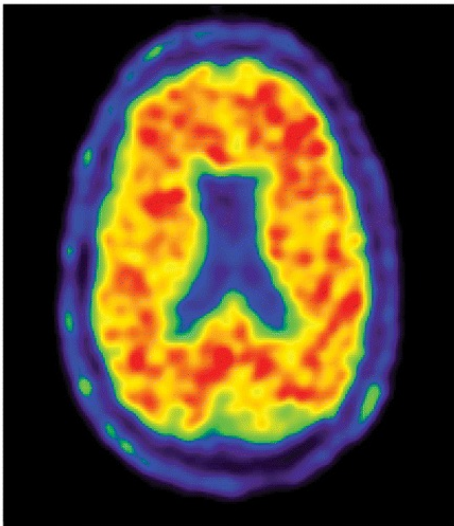
M. ten Kate et al., Alz Res Ther (2018)

See also: D. Cash et al., Alz Res Ther (2014)

NORMAL



ABNORMAL

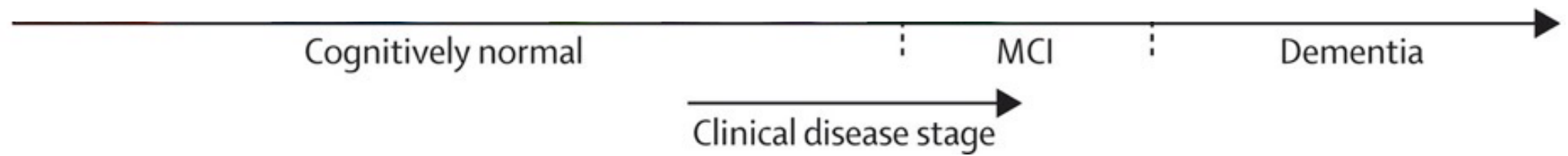


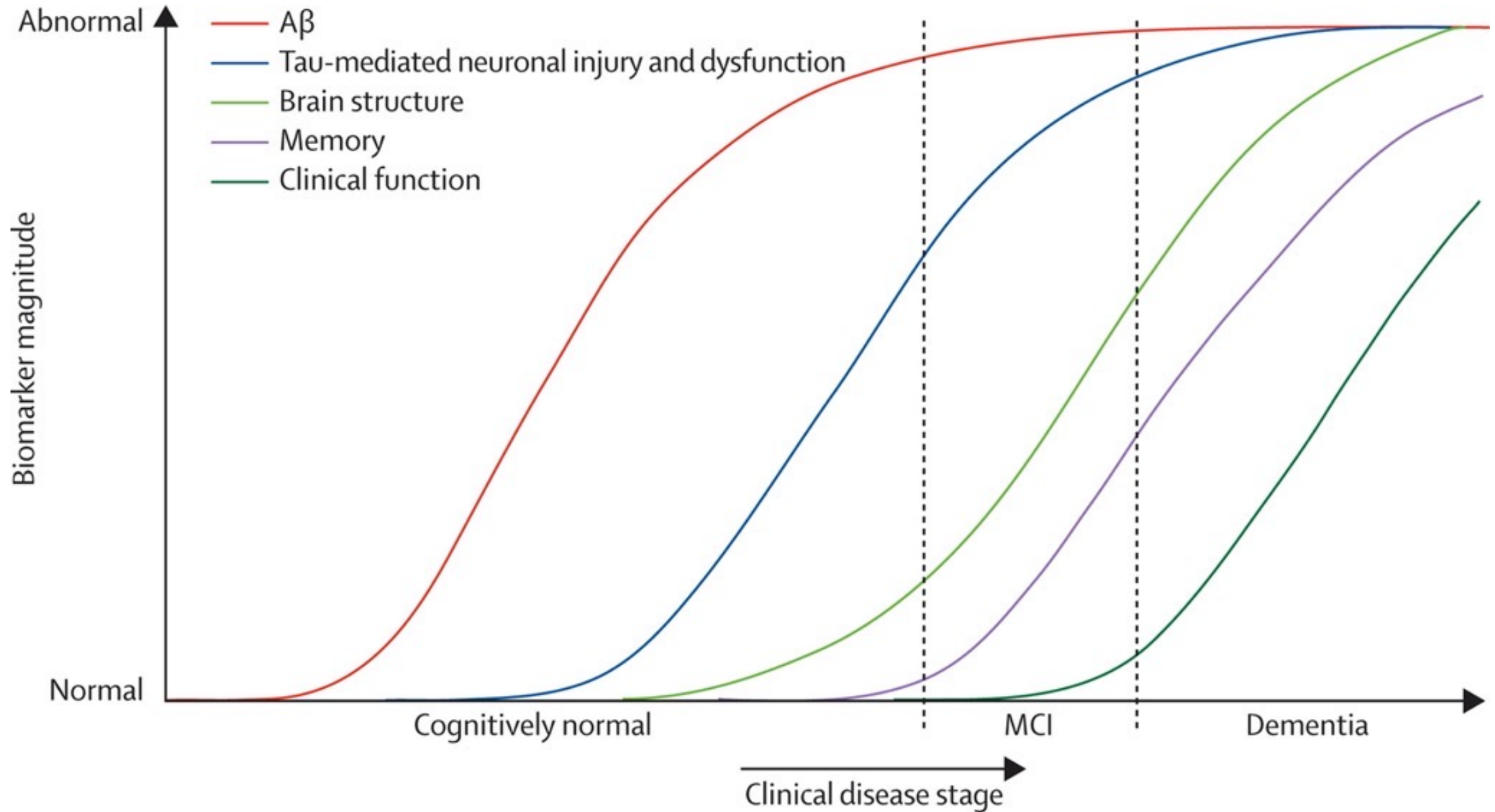
AMYLOID PET

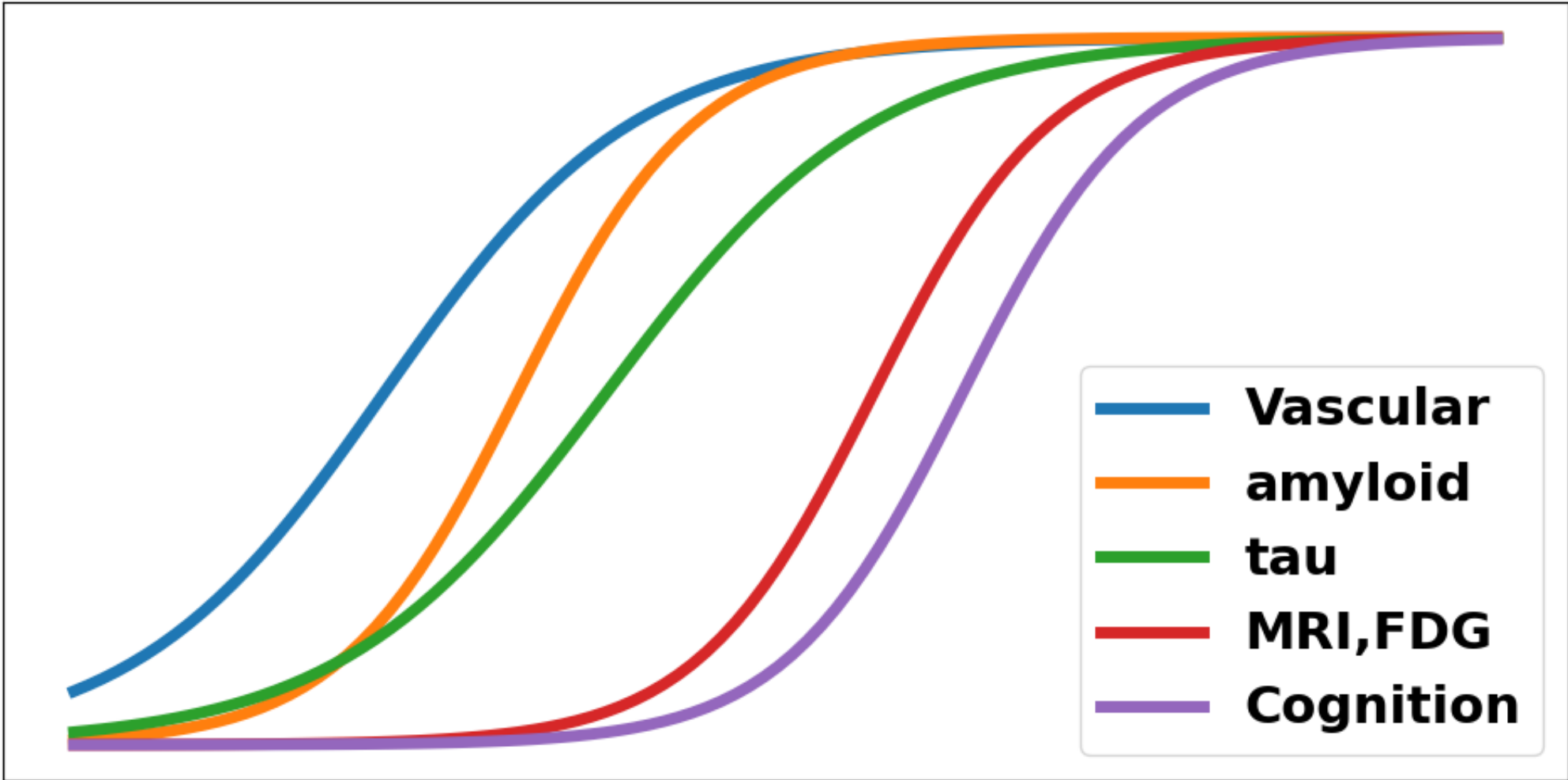
TAU PET

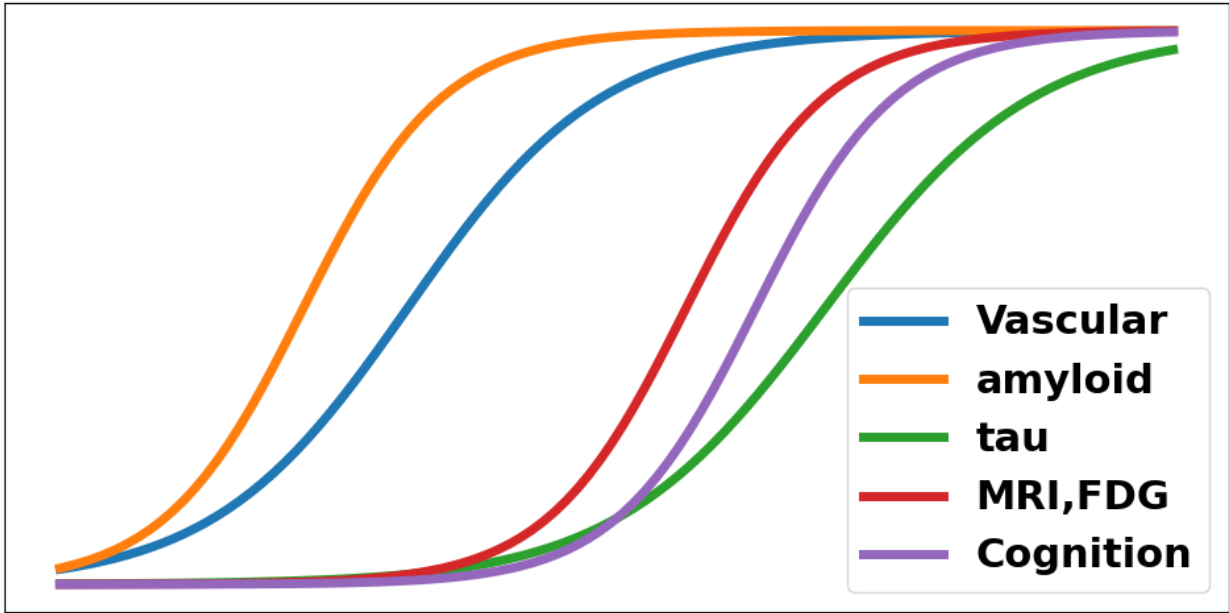
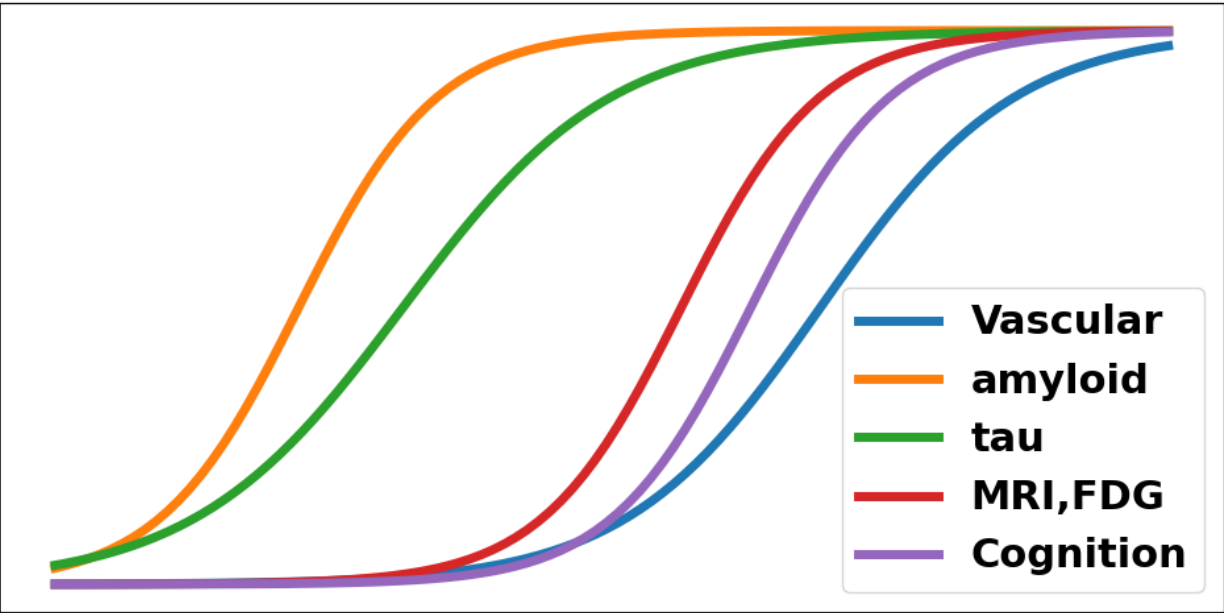
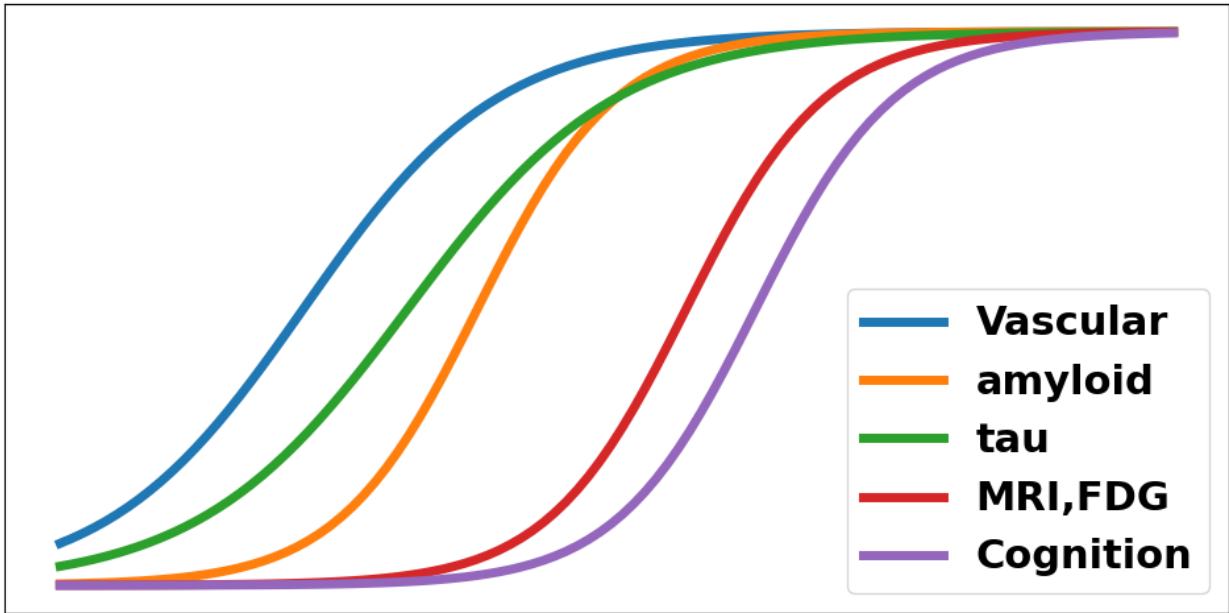
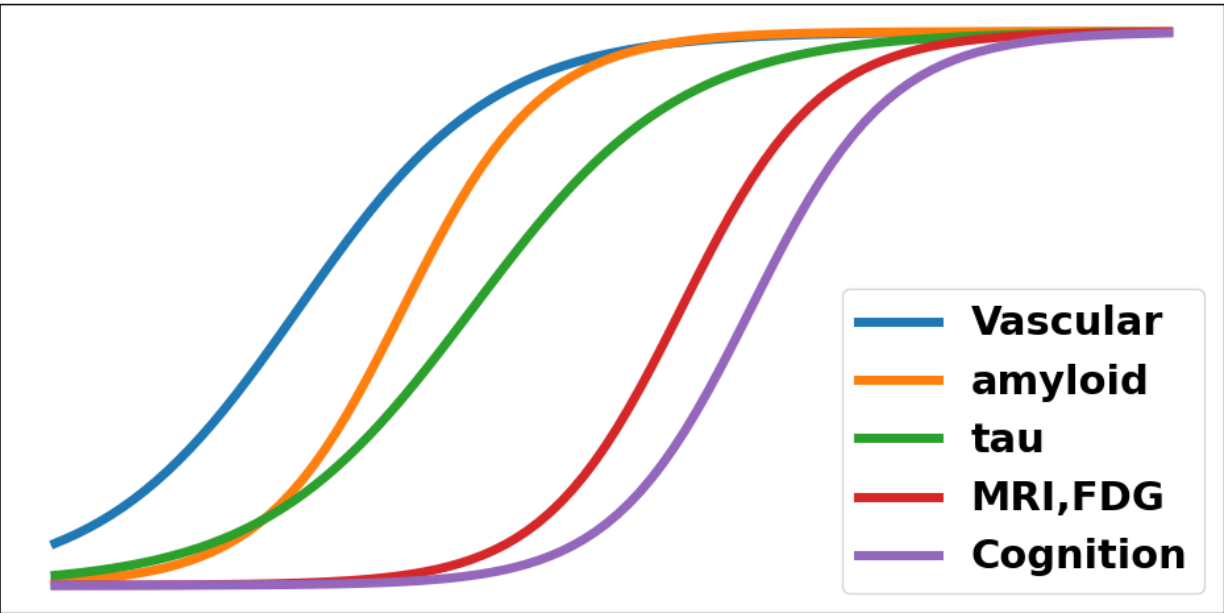
FDG PET

What have
clinical
trials
done?

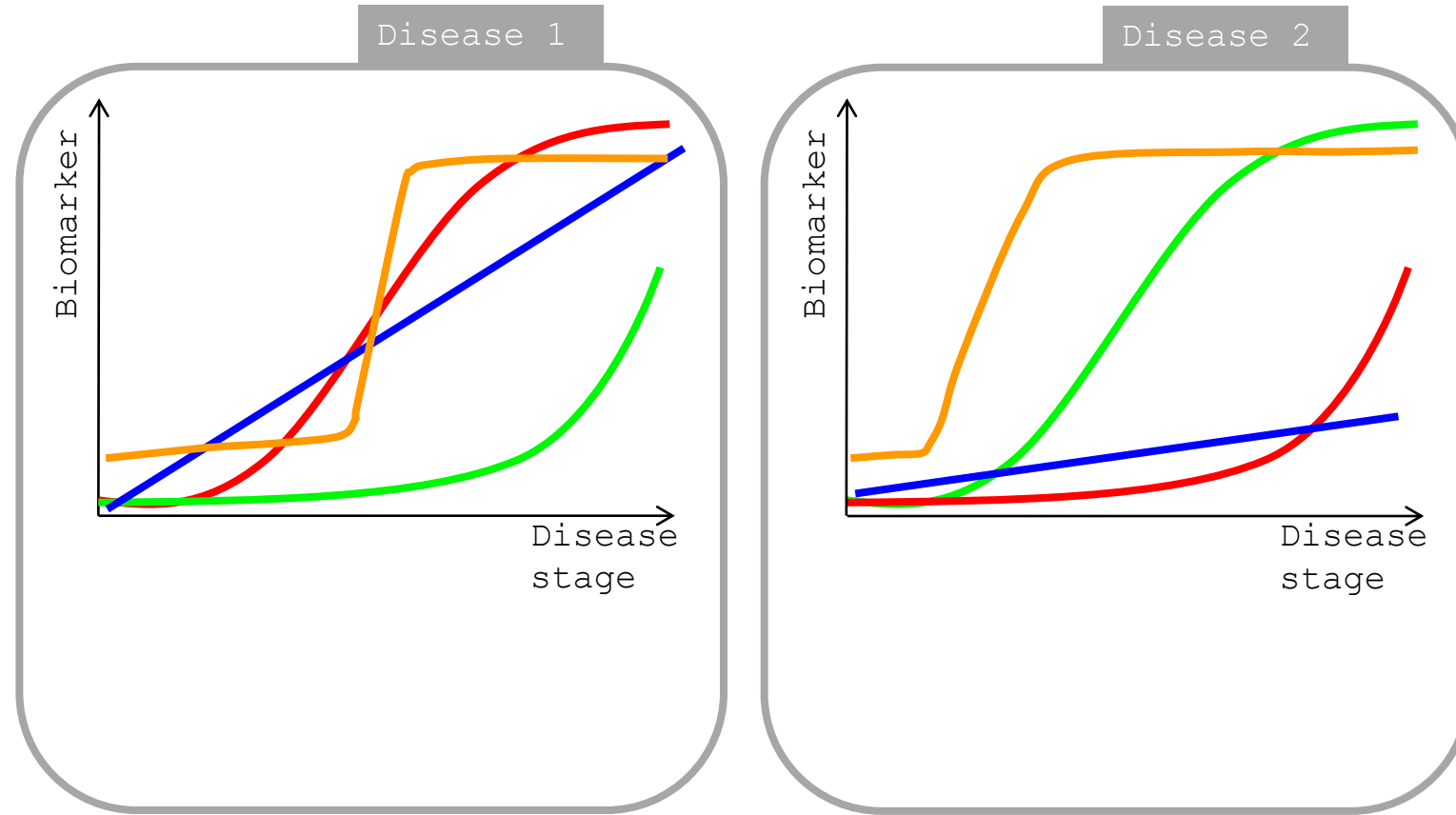




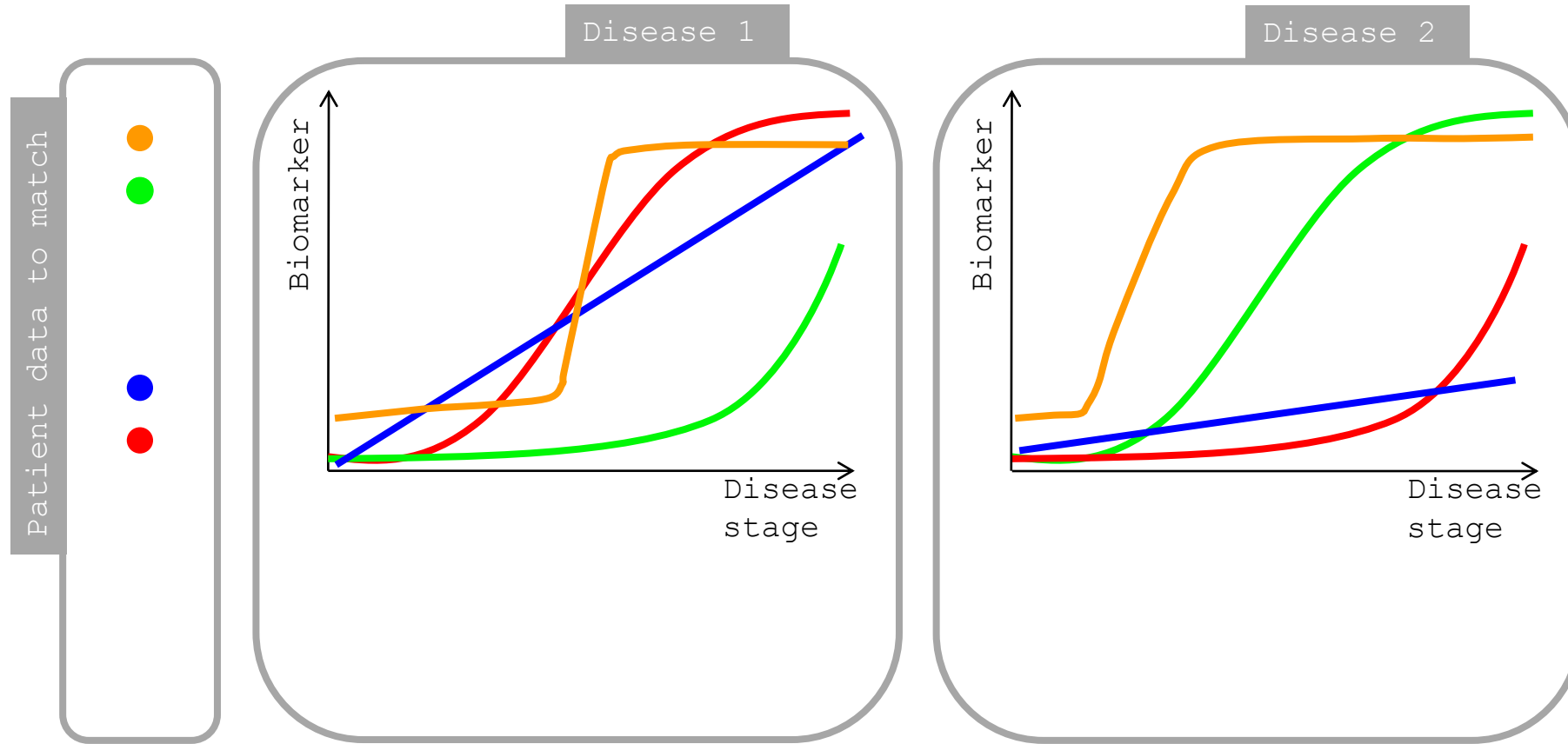




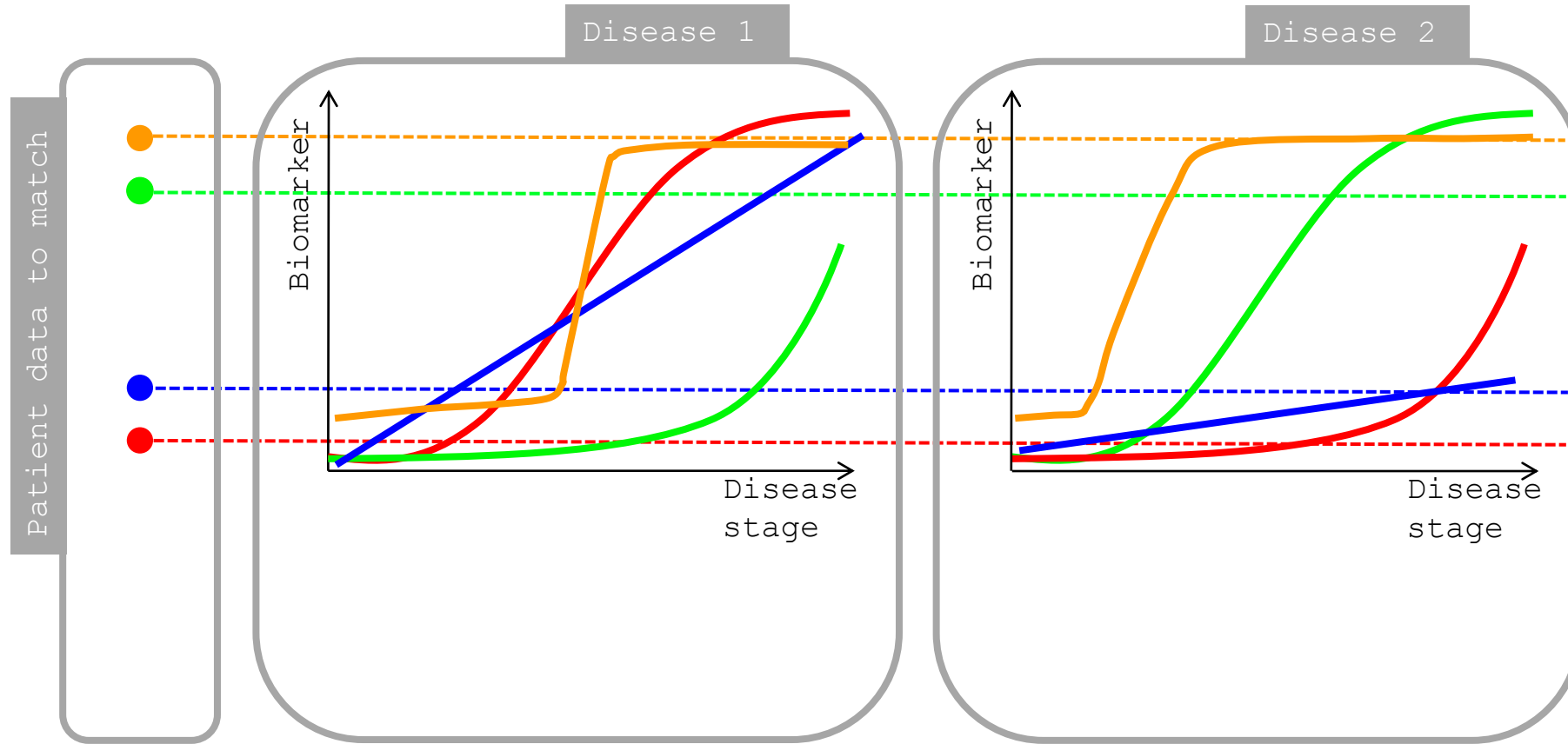
Quantitative utility of data-driven models



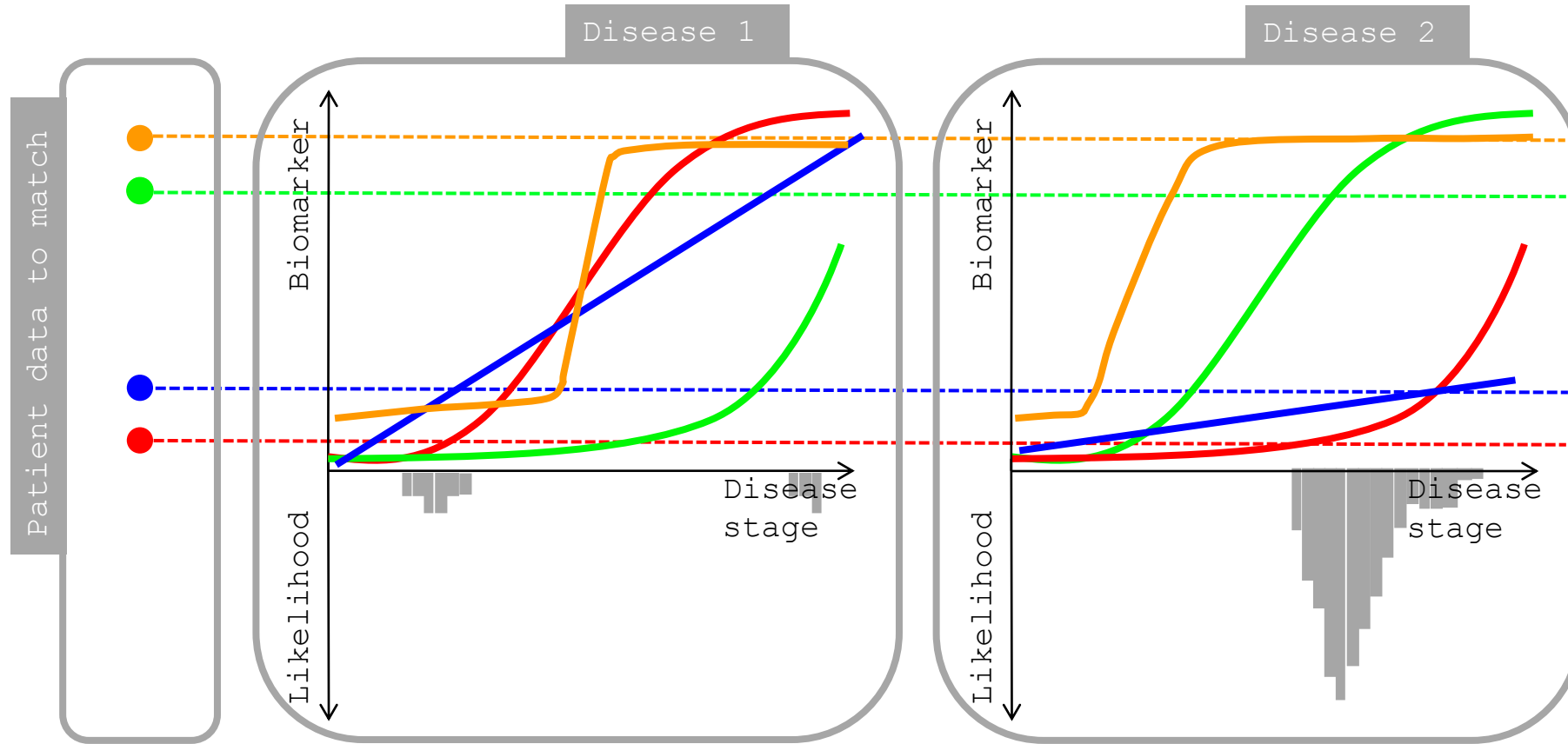
Quantitative utility of data-driven models



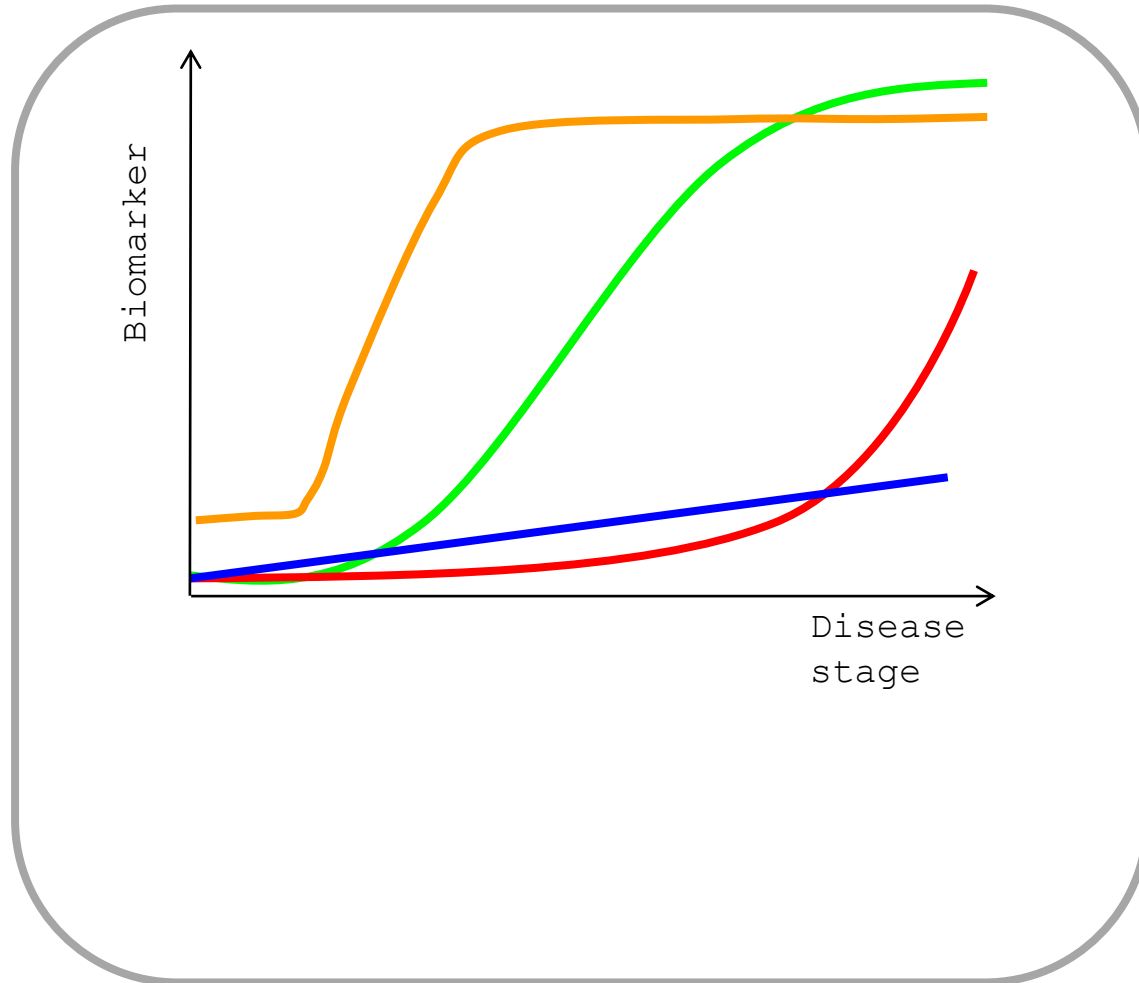
Quantitative utility of data-driven models



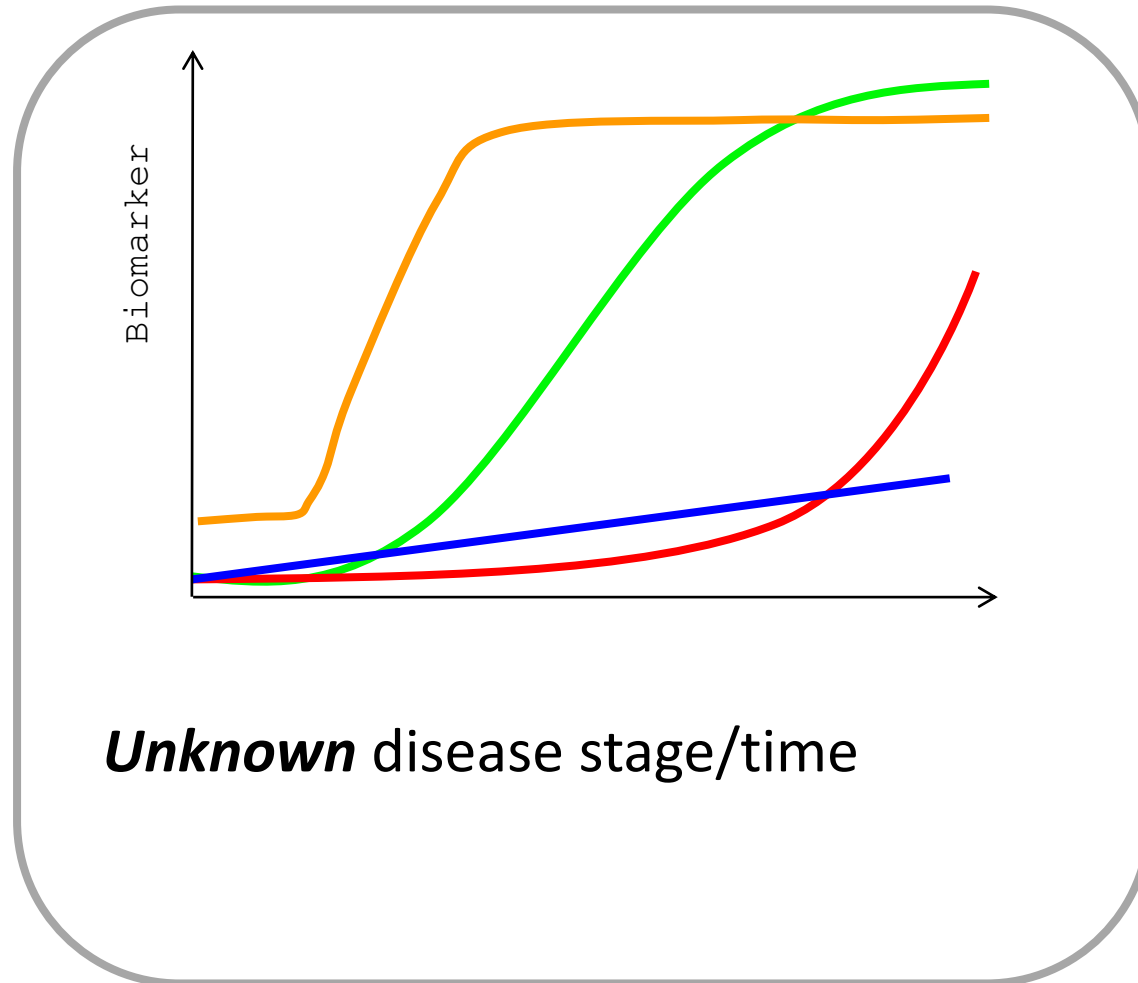
Quantitative utility of data-driven models



Traditional models



Traditional models



The Journey to Data-driven disease progression modelling

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

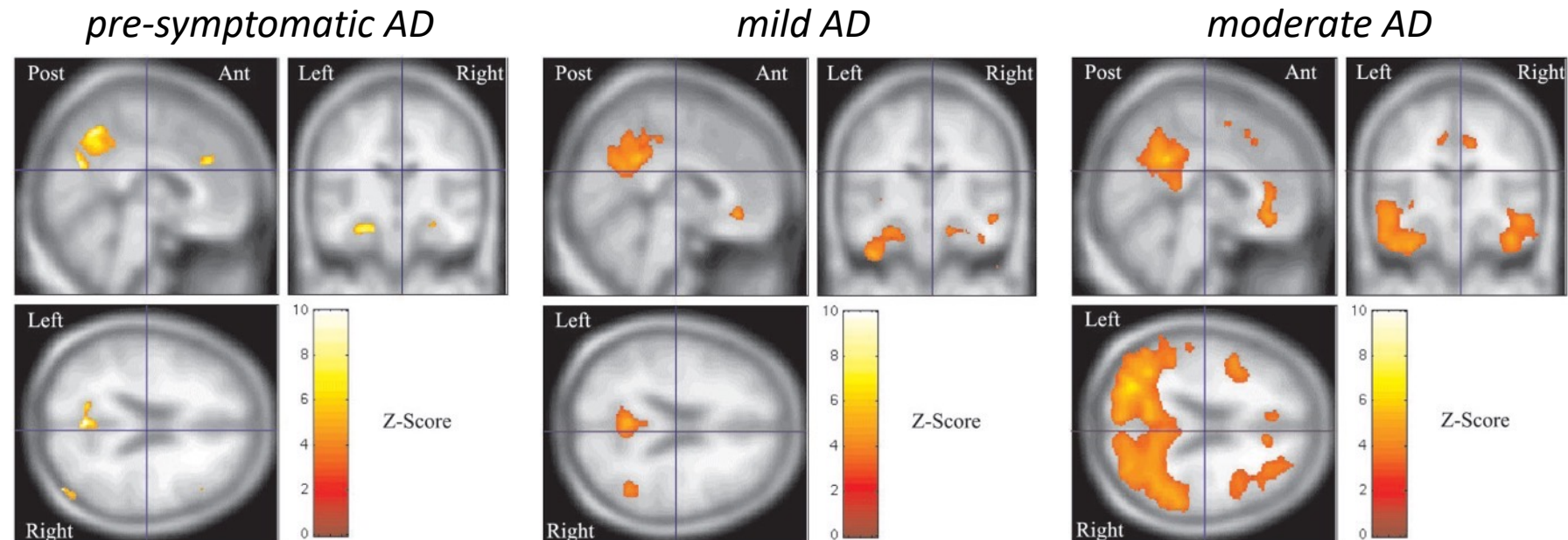
The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression

Scahill et al. PNAS 2002

- T1 MRI measures of neuronal atrophy: MMSE “clock”



The Journey to Data-driven disease progression modelling

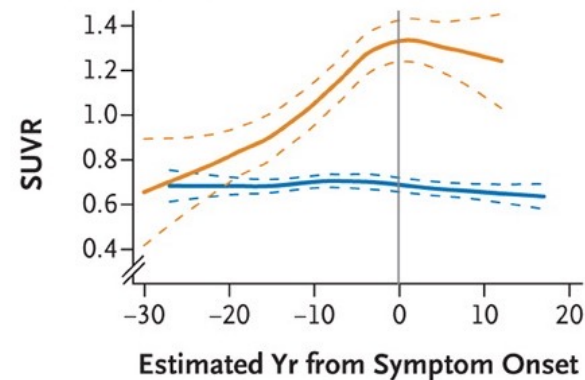
2002–2008 Traditional: stage == symptoms

- Regression

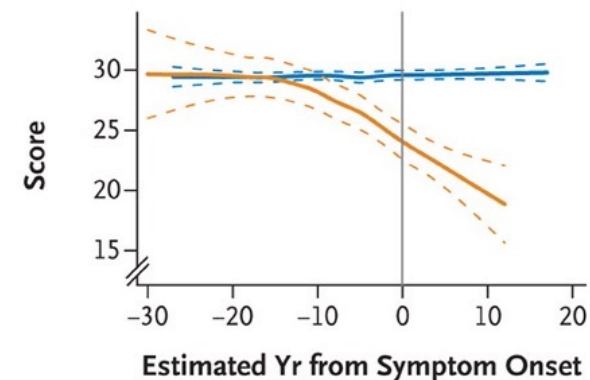
Bateman et al. NEJM 2012

- Parental age of symptom onset in dominantly-inherited AD

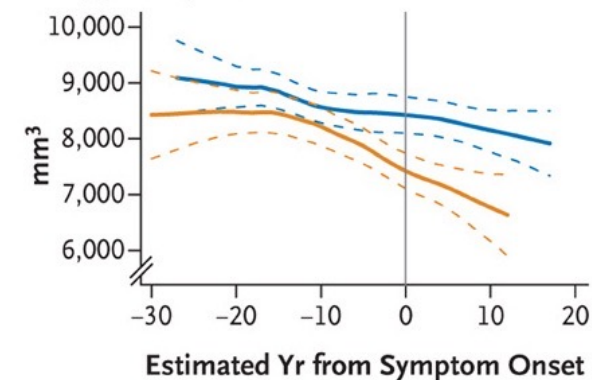
F $A\beta$ Deposition in the Precuneus



B Mini-Mental State Examination



D Hippocampal Volume

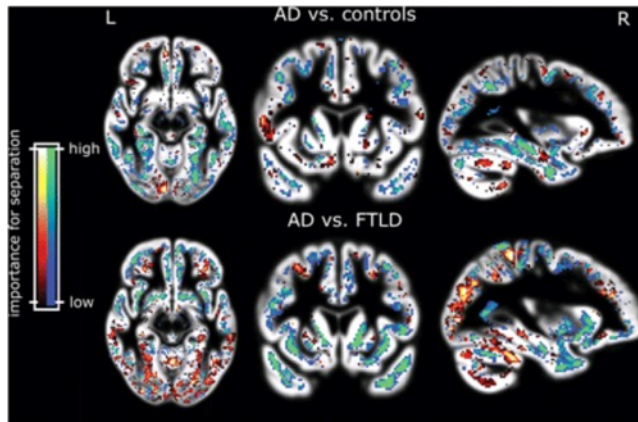


The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression
- Pattern recognition (supervised ML)

Classifying structural MRI in AD



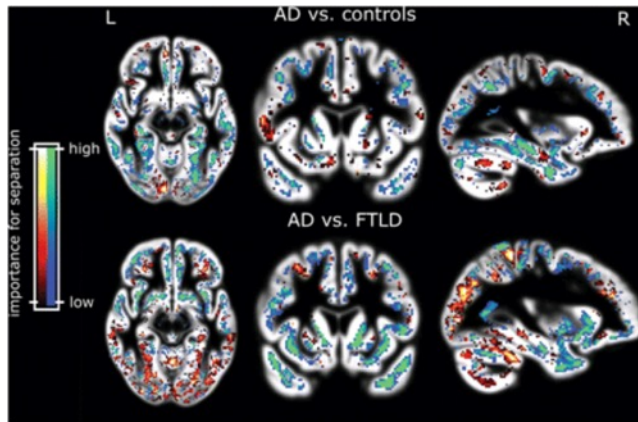
Klöppel et al. Brain 2008

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

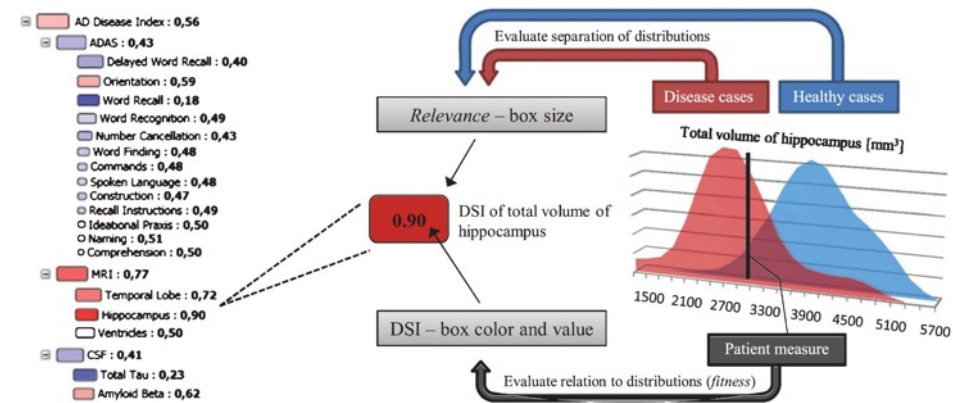
- Regression
- Pattern recognition (supervised ML)

Classifying structural MRI in AD



Klöppel et al. Brain 2008

Disease State Fingerprint for AD



Mattila et al. JAD 2011

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative



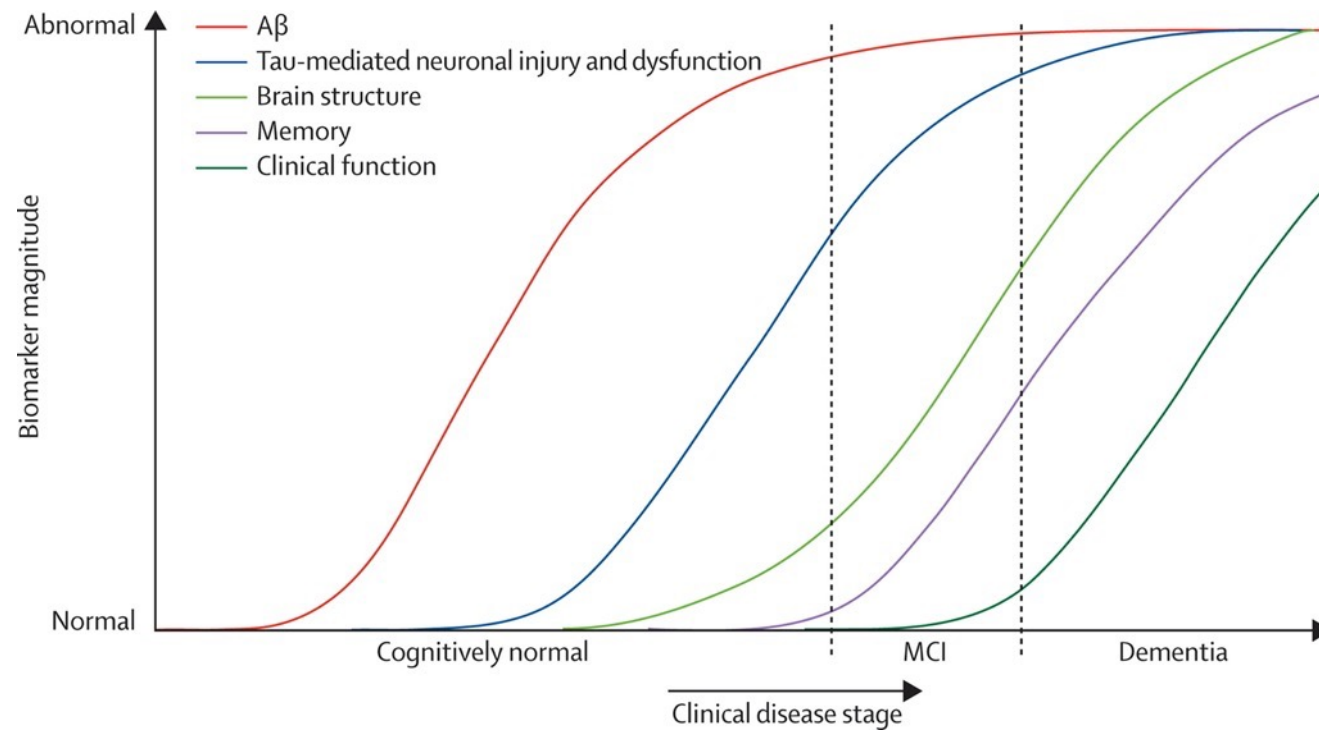
The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression



Jack et al. TLN 2010

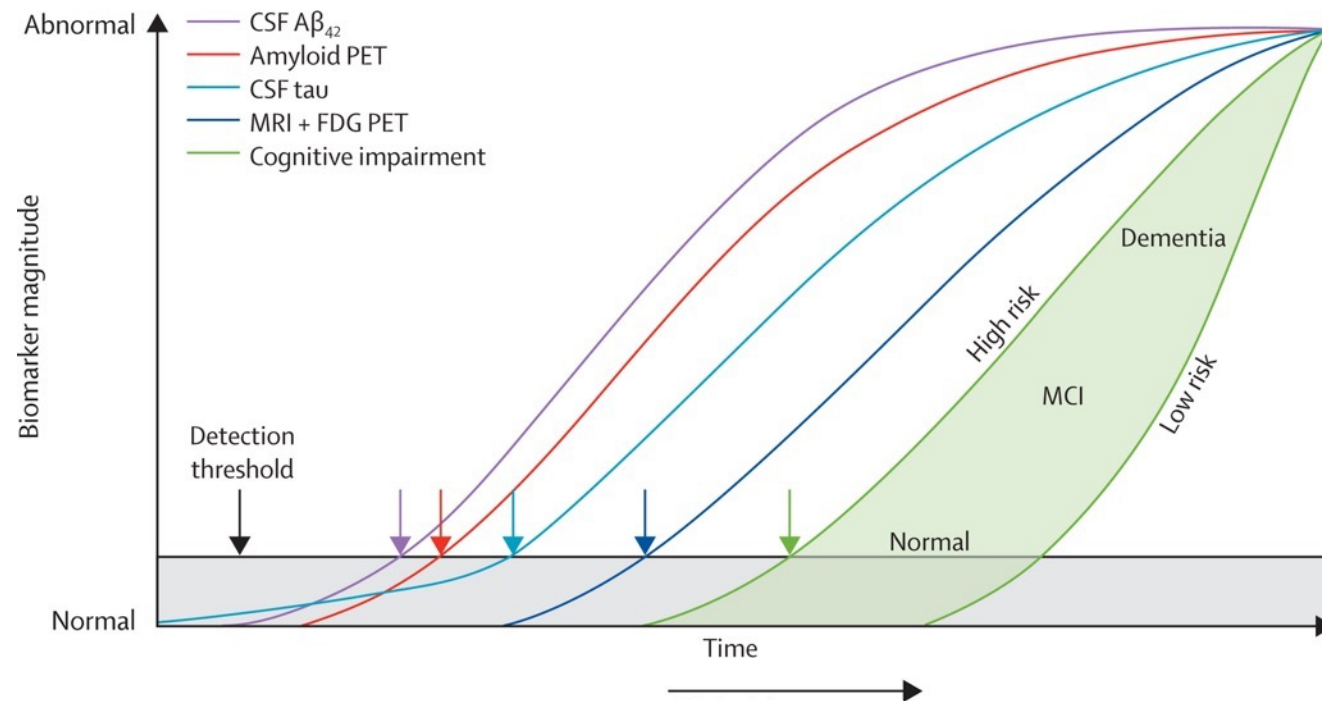
The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression (updated 2013)



Jack et al. TLN 2013

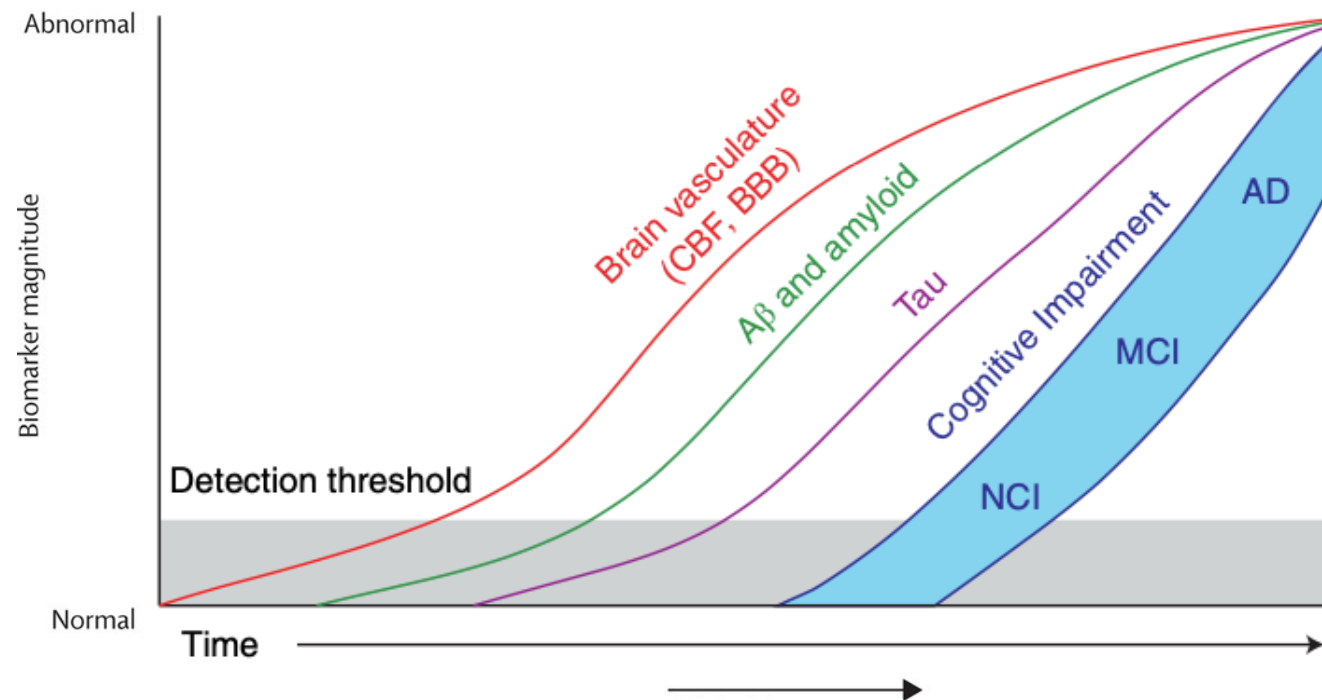
The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression (updated 2018)



Sweeney et al. Nat Comms
2018

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression

2011 Data-Driven Disease Progression Modelling

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression

2011 Data-Driven Disease Progression Modelling

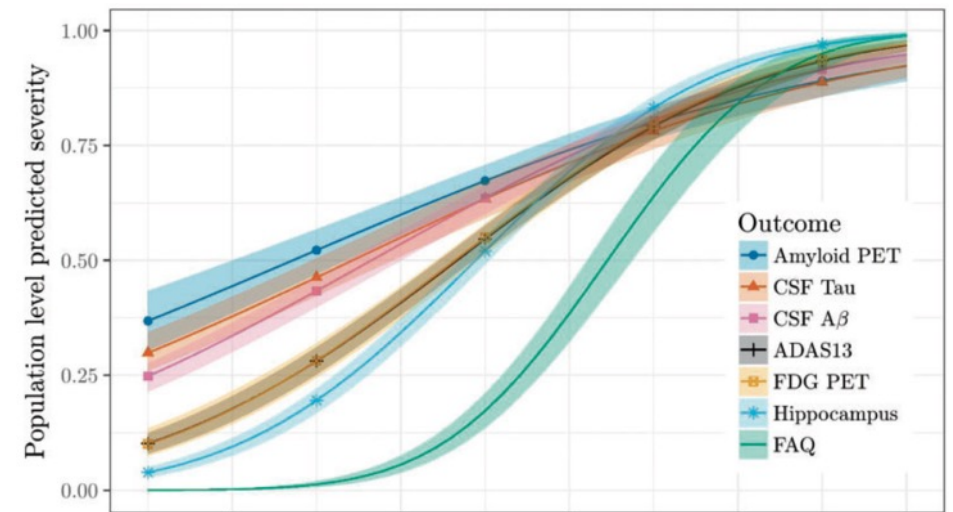
- Pseudo-time methods:
 - discrete (EBM sequencing)
 - continuous (latent-time: LTJMM, IRT, GPPM)

Li et al. Stat Meth Med Res 2017 (2014)

Leoutsakos et al. JPAD 2016

Schiratti et al. MICCAI 2015, JMLR 2017

Lorenzi et al. NIMG 2017



The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progress

2011 Data-Driven Disease Progression Modelling

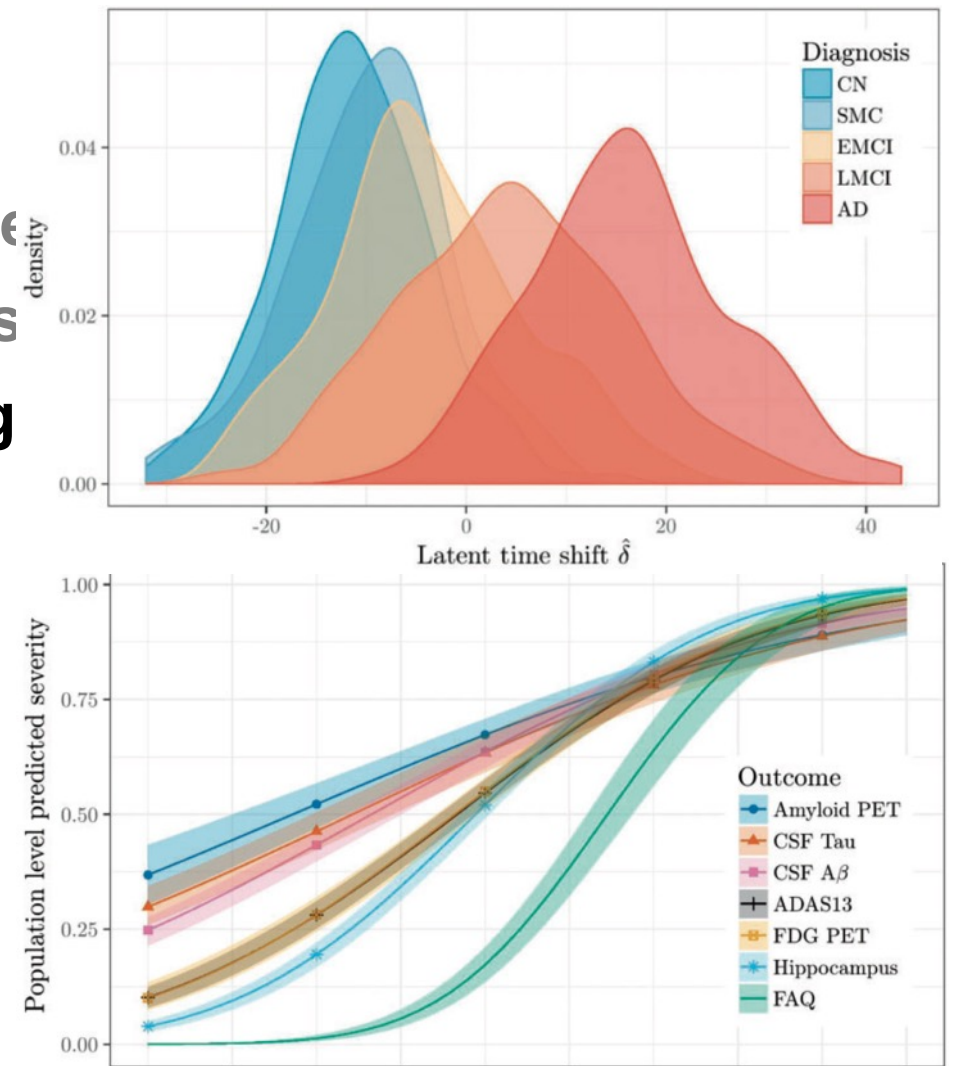
- Pseudo-time methods:
 - discrete (EBM sequencing)
 - continuous (latent-time: LTJMM, IRT, GPPM)

Li et al. Stat Meth Med Res 2017 (2014)

Leoutsakos et al. JPAD 2016

Schiratti et al. MICCAI 2015, JMLR 2017

Lorenzi et al. NIMG 2017



The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

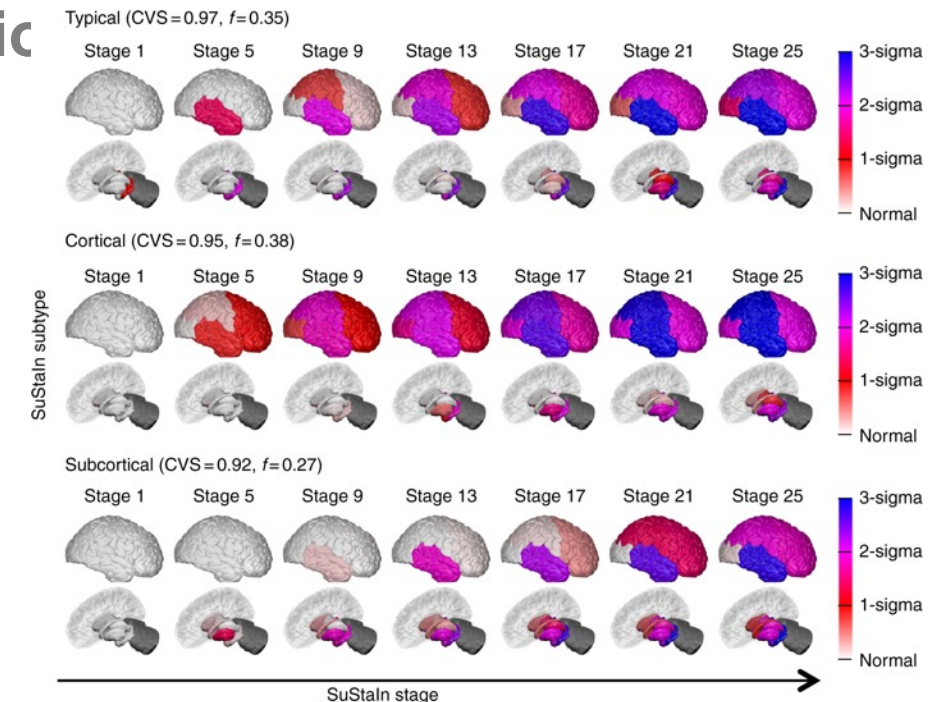
2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression

2011 Data-Driven Disease Progression Modelling

- Pseudo-time methods:
 - discrete (EBM sequencing)
 - continuous (latent-time: LTJMM, IRT, GPPM)
- Pseudo-time + Clustering

Subtype & Stage Inference (SuStaln)



Young et al. Nat. Comms 2018

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

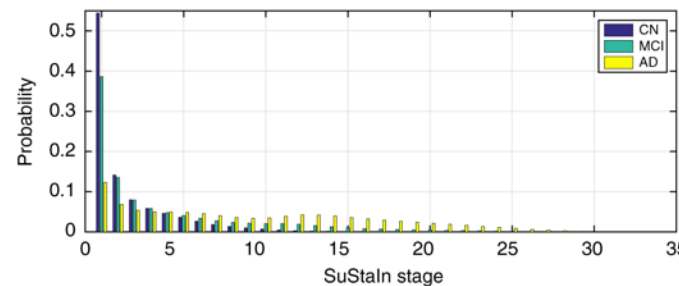
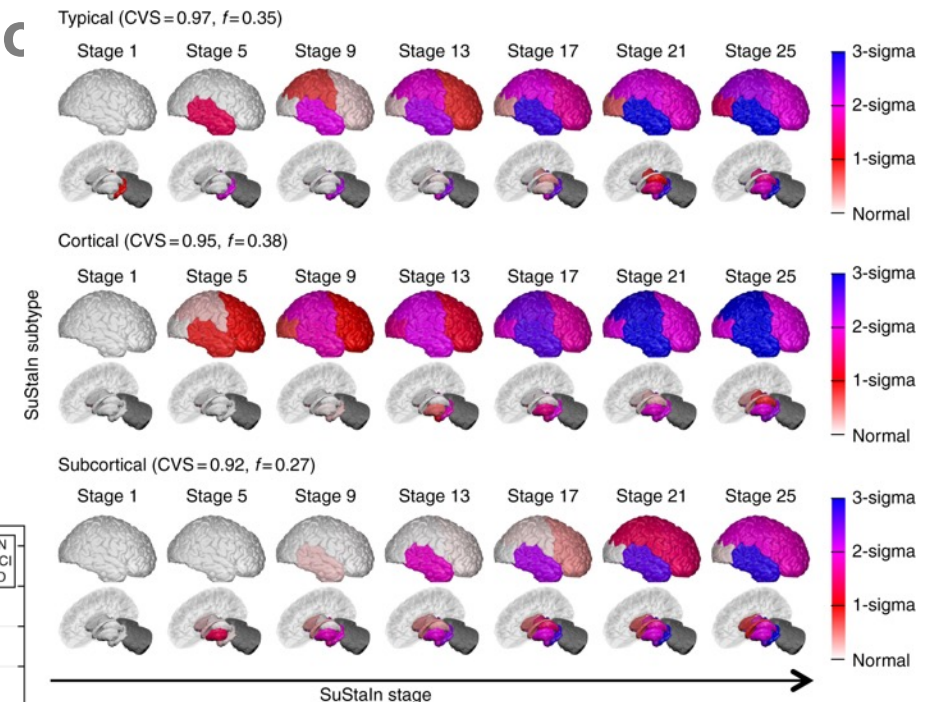
2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression

2011 Data-Driven Disease Progression Modelling

- Pseudo-time methods:
 - discrete (EBM sequencing)
 - continuous (latent-time: LTJMM, IRT, GPPM)
- Pseudo-time + Clustering

Subtype & Stage Inference (SuStaln)



Young et al. Nat. Comms 2018

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

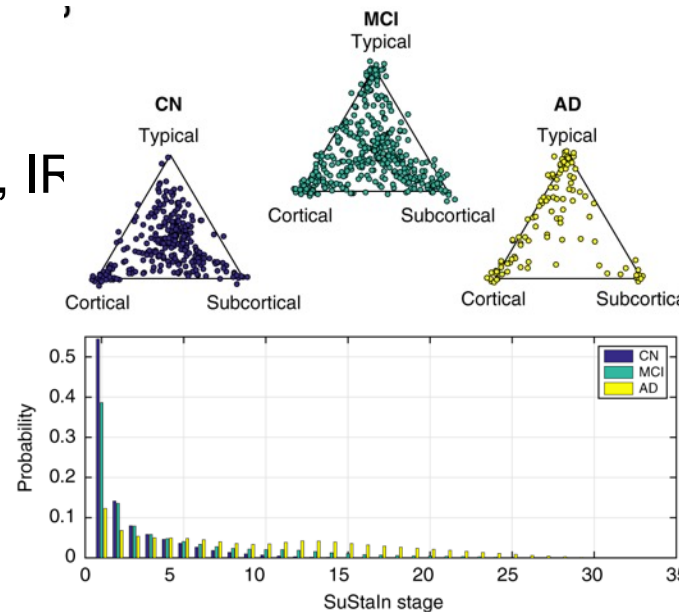
- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

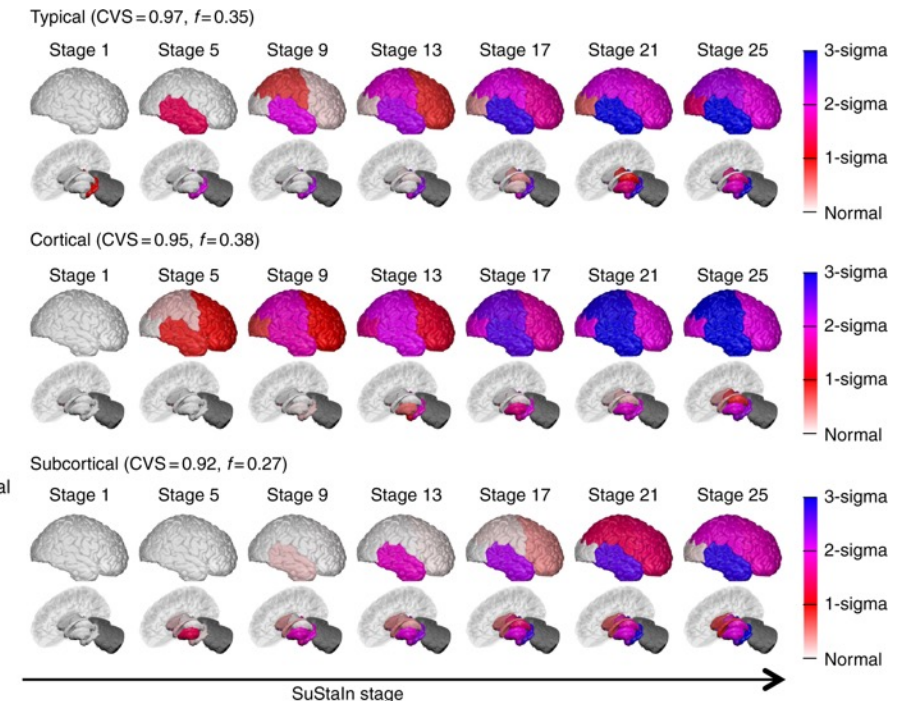
2010 Hypothetical Models of Alzheimer's progression

2011 Data-Driven Disease Progression Modelling

- Pseudo-time methods:
 - discrete (EBM sequencing)
 - continuous (latent-time: LTJMM, IF)
- Pseudo-time + Clustering



Subtype & Stage Inference (SuStaln)



Young et al. Nat. Comms 2018

The Journey to Data-driven disease progression modelling

2002–2008 Traditional: stage == symptoms

- Regression, Pattern recognition (supervised ML)

2004 Alzheimer's Disease Neuroimaging Initiative

2010 Hypothetical Models of Alzheimer's progression

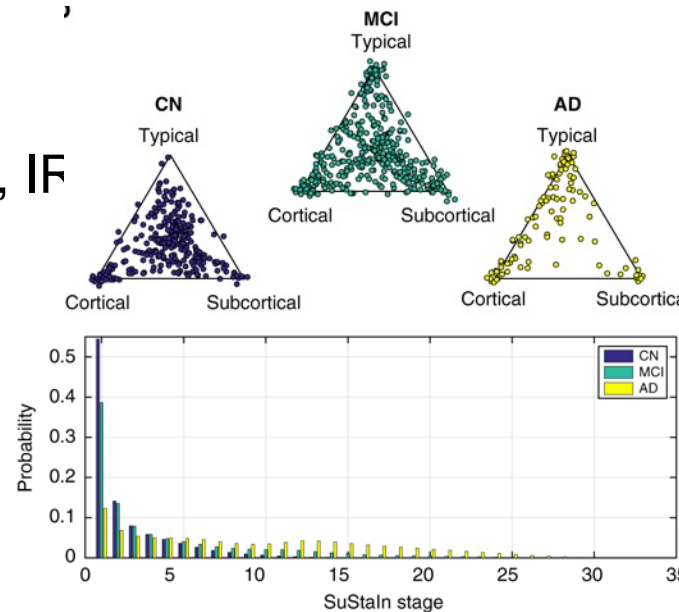
2011 Data-Driven Disease Progression Modelling

- Pseudo-time methods:
 - discrete (EBM sequencing)
 - continuous (latent-time: LTJMM, IF)
- Pseudo-time + Clustering

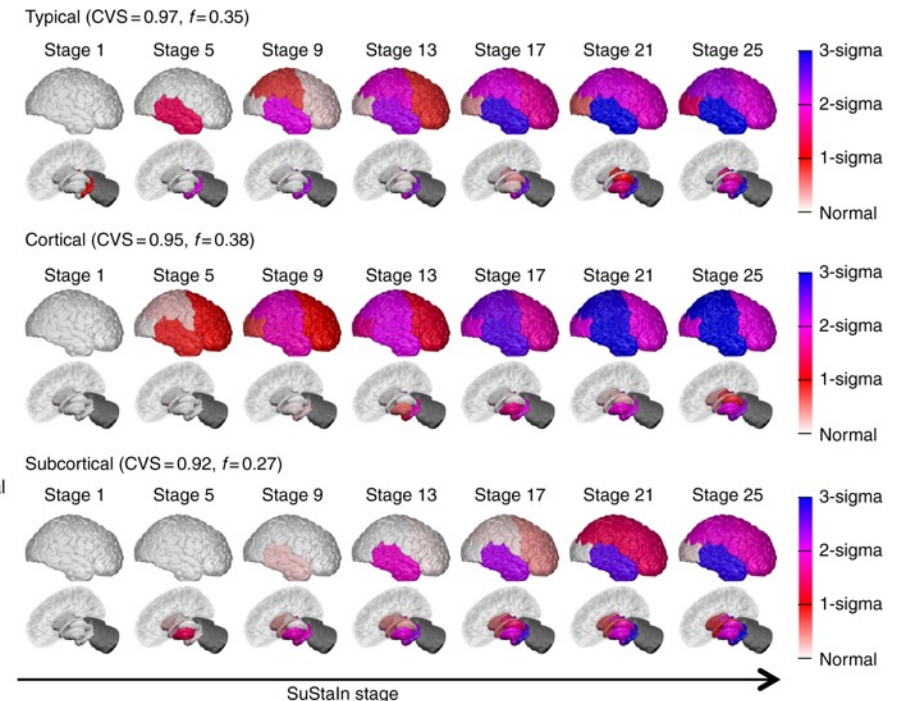
• *tau* PET: Vogel+ Nat Med 2021

• *Amyloid+tau*:

Aksman bioRxiv 2020, Brain 2023



Subtype & Stage Inference (SuStaln)



Young et al. Nat. Comms 2018

Data-driven disease progression models

So far:

Phenomenological models

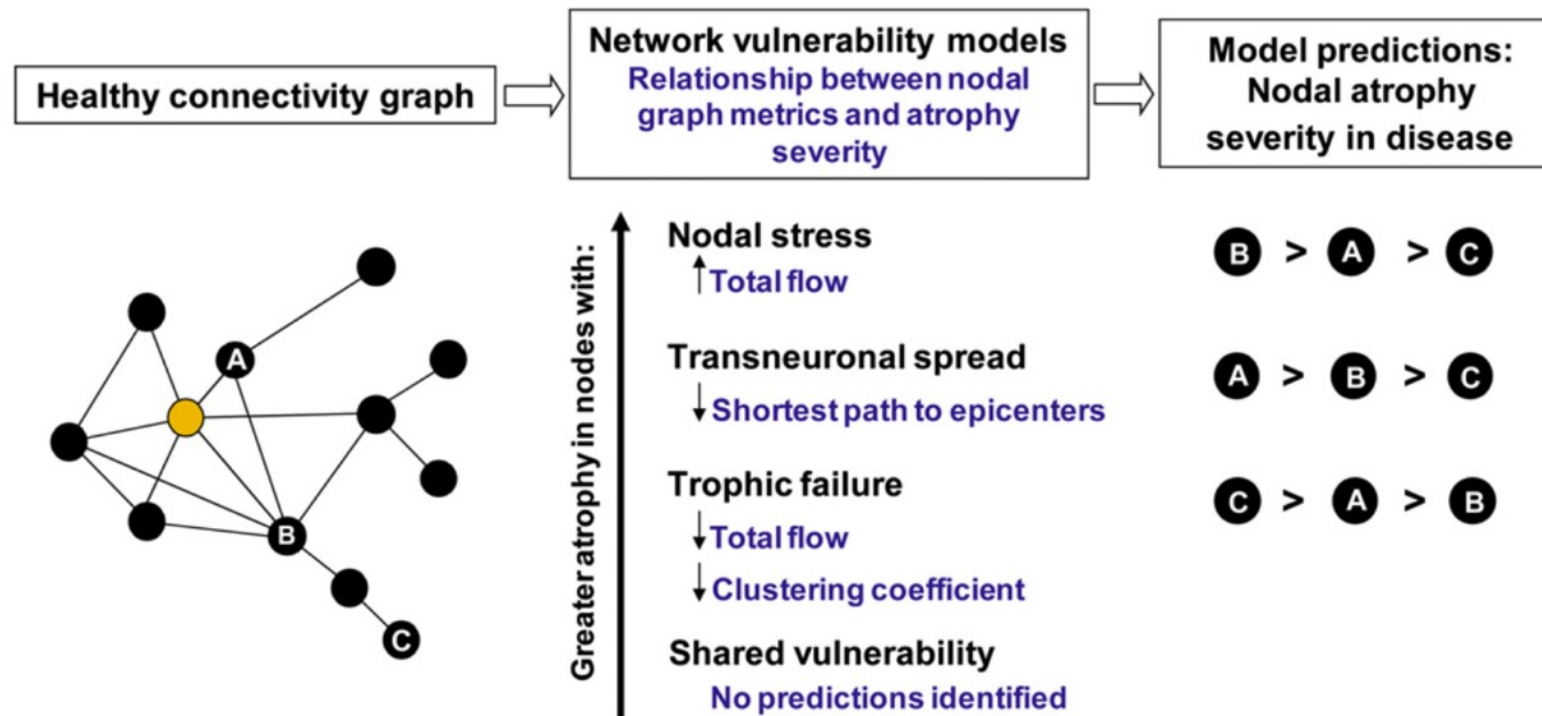
What about disease *mechanisms*?

Can we understand/explain
“Top-down” observations of pathology, using
“Bottom-up” models of mechanism/physiology?

Bottom-up models

2009–2012 Hypotheses of neurodegeneration due to pathogens

- Selective vulnerability / Wear-and-tear / Network / Use-it-or-lose-it
- Seeley et al. Neuron 2009, Zhou et al. Neuron 2012



Bottom-up models

2009–2012 Hypotheses of neurodegeneration

- Seeley et al. Neuron 2009, Zhou et al. 2012

2012– Protein (prion) Spreading Models

- **2012:** Network diffusion model (heat eq)
- **2014:** Epidemic Spreading Model

Raj et al. Neuron 2012

Iturria-Medina+ PLOS Comp. Biol. 2014

Bottom-up models

2009–2012 Hypotheses of neurodege

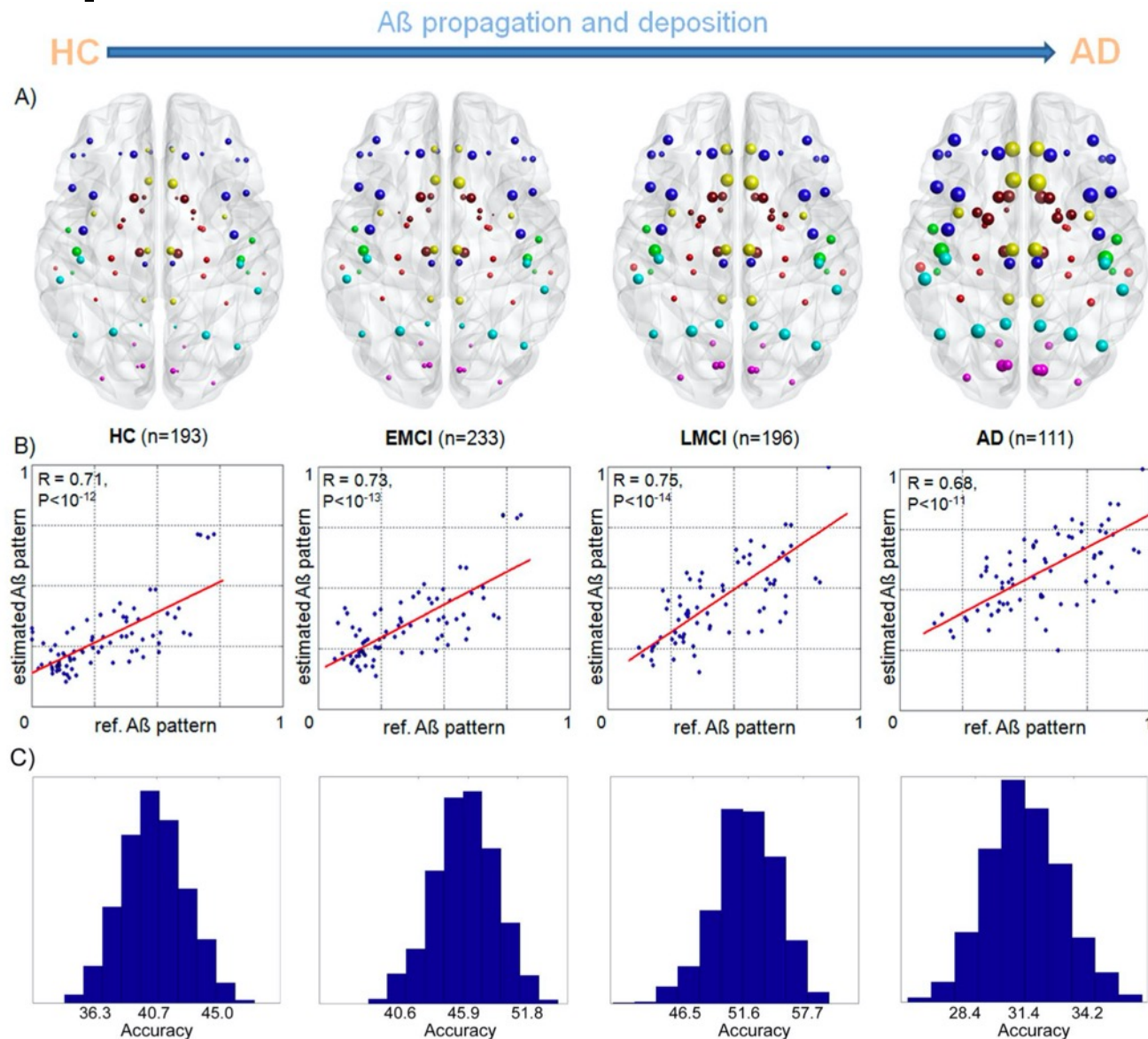
- Seeley et al. Neuron 2009, Zhou et

2012– Protein (prion) Spreading Model

- 2012: Network diffusion model (heat eq)
- 2014: Epidemic Spreading Model

Raj et al. Neuron 2012

Iturria-Medina+ PLOS Comp. Biol. 2014



Bottom-up models

2009–2012 Hypotheses of neurodegeneration

- Seeley et al. Neuron 2009, Zhou et al. 2012

2012– Protein (prion) Spreading Models

- **2012:** Network diffusion model (heat eq) Multiphysics of Prionlike Diseases: Progression and Atrophy

- **2014:** Epidemic Spreading Model

Johannes Weickenmeier, Ellen Kuhl, and Alain Goriely
Phys. Rev. Lett. **121**, 158101 – Published 12 October 2018

 See Focus story: [A Physical Model for Neurodegenerative Disease](#)

- **2018–19: Physics** (Network Spreading + misfolding kinetics)

- Fisher-Kolmogorov = reaction-diffusion eq. (no mechanistic insight)
- Heterodimer = normal & abnormal proteins (+ clearance/production)
- Smoluchowski = stat. physics workhorse (+ size of protein aggregates)

Weickenmeier et al. Phys Rev Lett 2018

Fornari et al. J.R.Soc. Interface 2019

Bottom-up models

2009–2012 Hypotheses of m

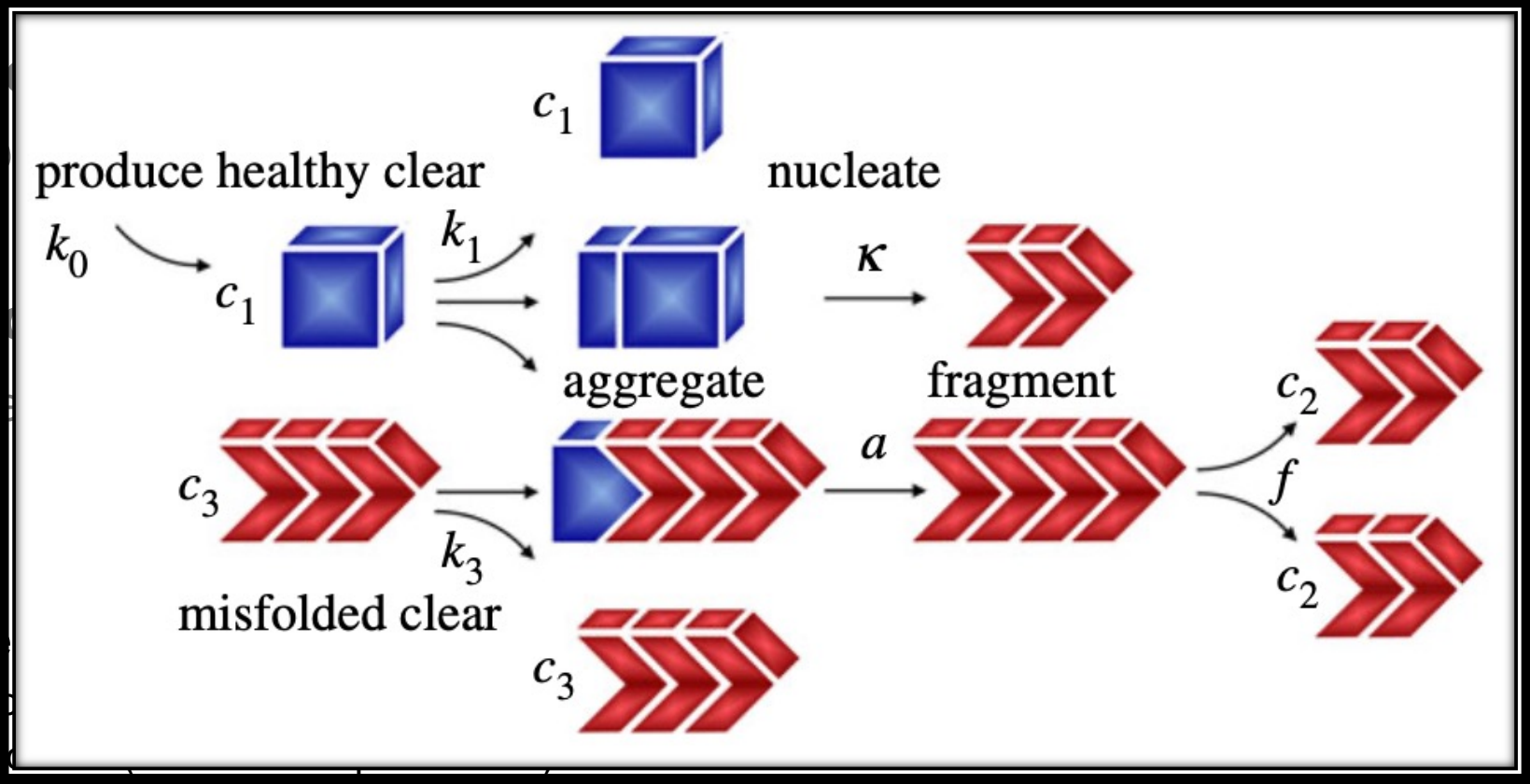
- Seeley et al. Neuron 2009

2012– Protein (prion) Spread

- 2012: Network diffusion model (h)
- 2014: Epidemic Spreading Model

2018–19: Physics (Network Spre

- Fisher-Kolmogorov = reaction-diffusion
- Heterodimer = normal & abnormal pr
- Smoluchowski = stat. physics workhorse (+ size of protein aggregates)



Weickenmeier et al. Phys Rev Lett 2018
 Fornari et al. J.R.Soc. Interface 2019

2023 tau *Production vs Transport* In Alzheimer's



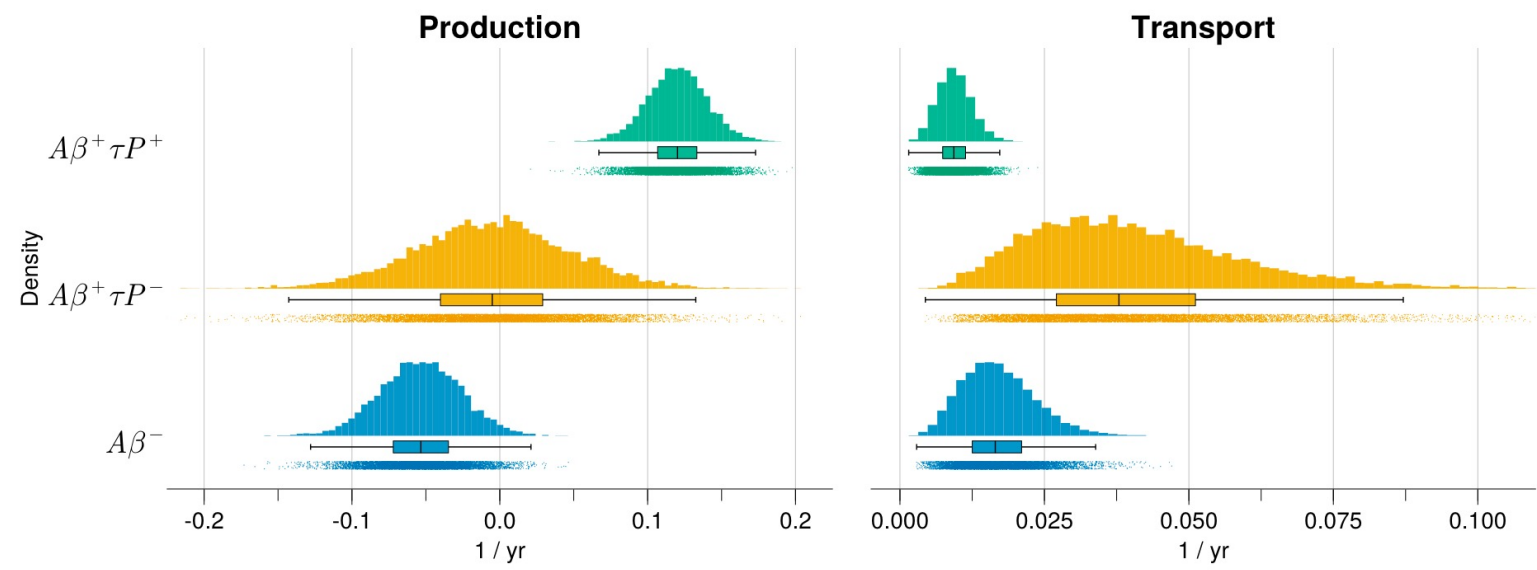
New Results

[Follow this preprint](#)

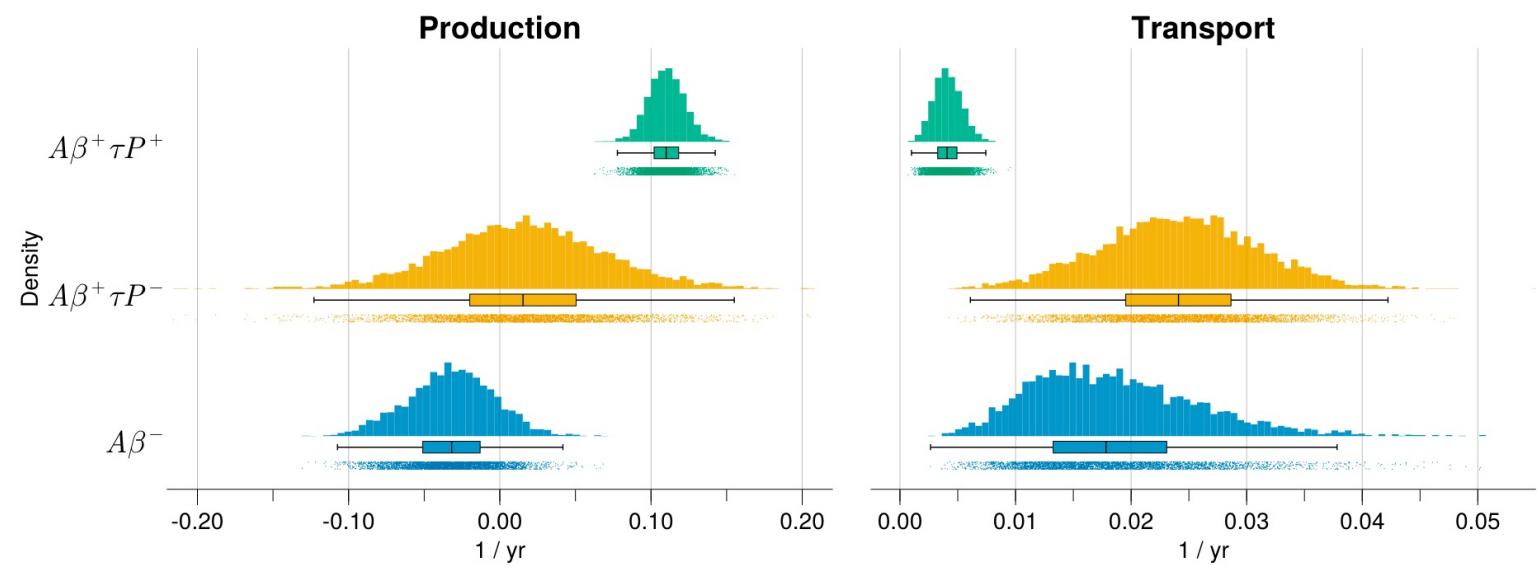
Personalised Regional Modelling Predicts Tau Progression in the Human Brain

[Pavanjit Chaggar](#), [Jacob Vogel](#), [Alexa Pichet Binette](#), [Travis B. Thompson](#), [Olof Strandberg](#),
[Niklas Mattsson-Carlgren](#), [Linda Karlsson](#), [Erik Stomrud](#), [Saad Jbabdi](#), [Stefano Magon](#), [Gregory Klein](#),
 the Alzheimer's Disease Neuroimaging Initiative, [Oskar Hansson](#), [Alain Goriely](#)

doi: <https://doi.org/10.1101/2023.09.28.559911>



(a)



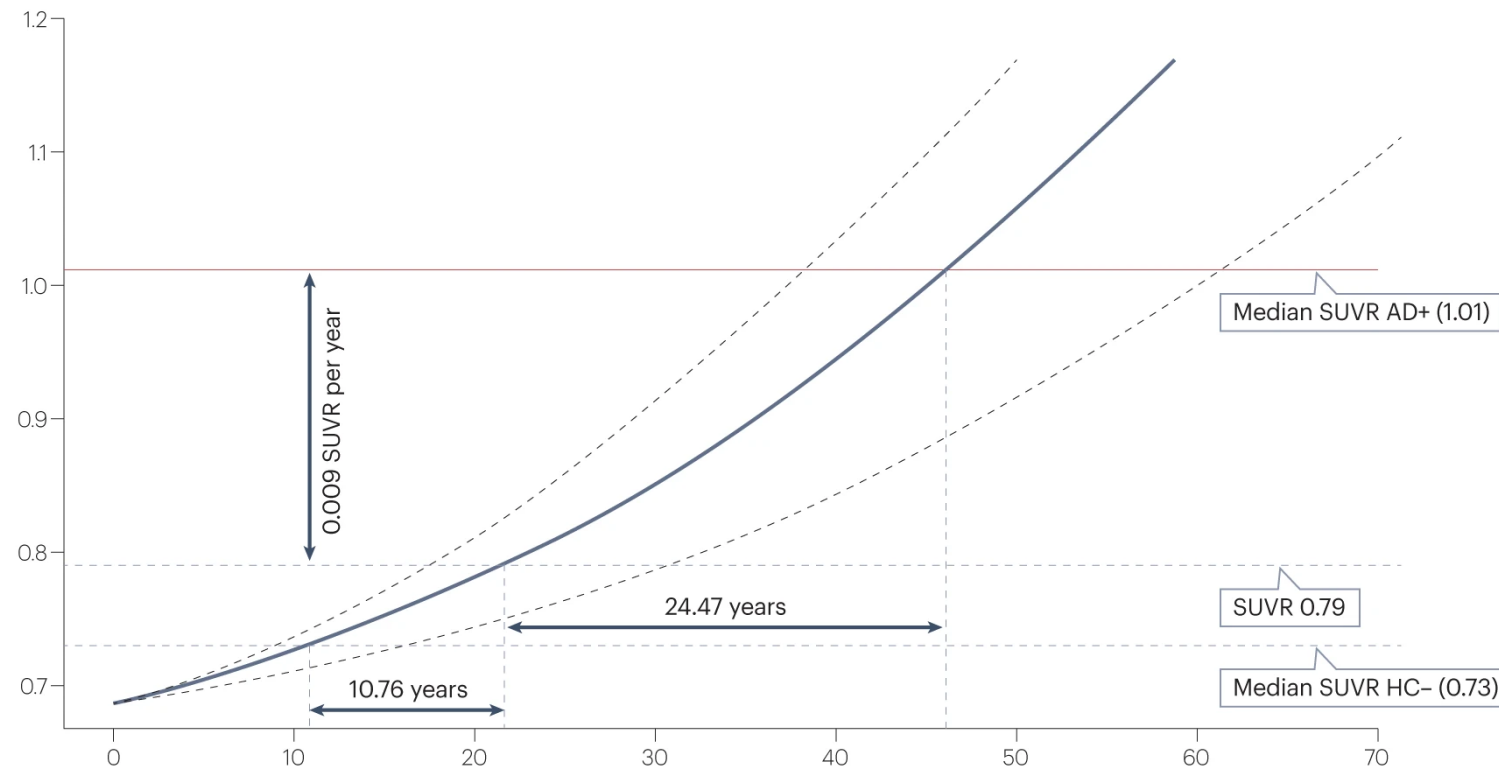
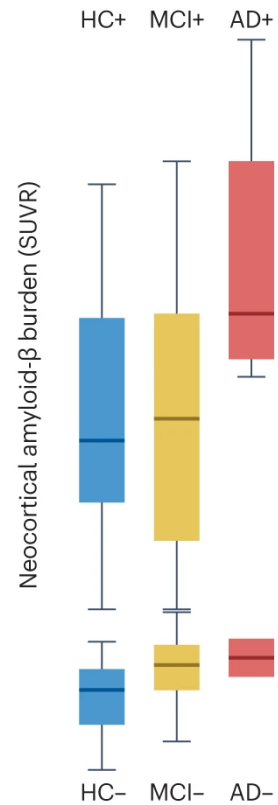
(b)

Applications

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young^{1,2,6}, Neil P. Oxtoby^{1,6}, Sara Garbarino³, Nick C. Fox⁴, Frederik Barkhof^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander¹

a Biological insight

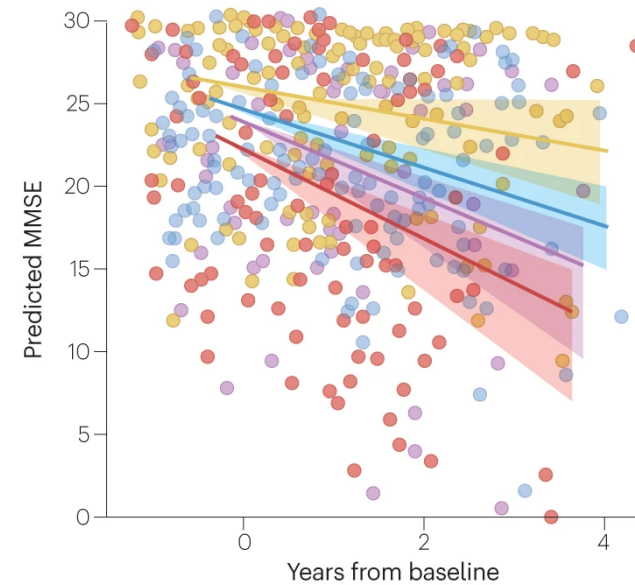
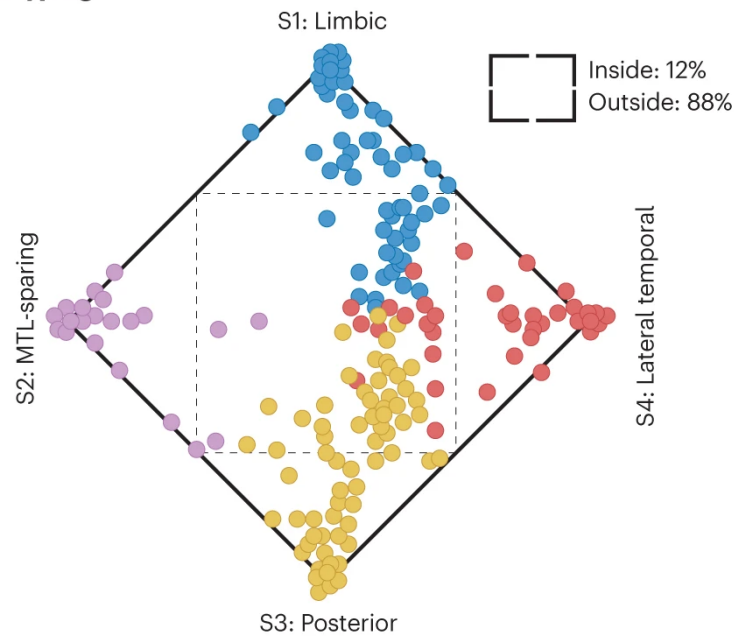


Applications

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young^{1,2,6}, Neil P. Oxtoby^{1,6}, Sara Garbarino³, Nick C. Fox⁴, Frederik Barkhof^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander¹

b Subtyping



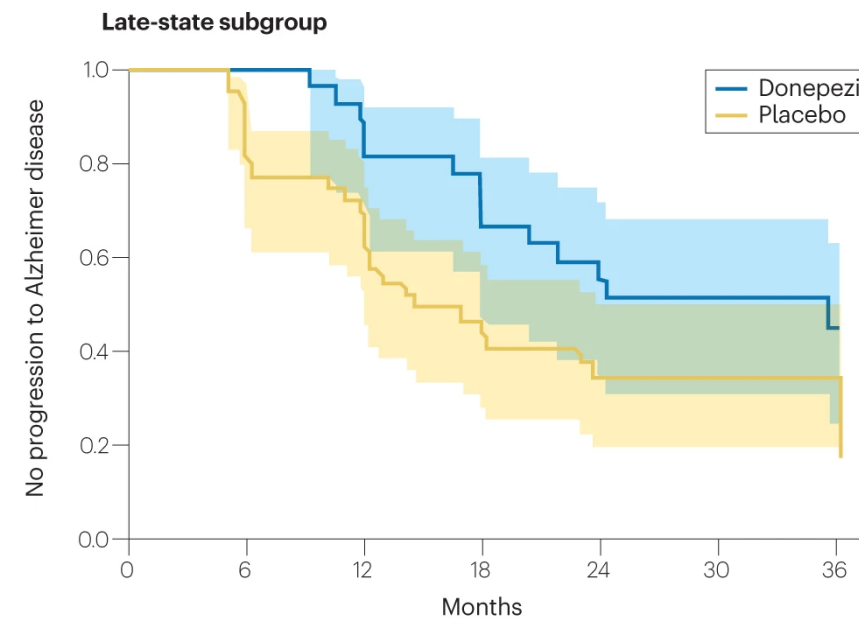
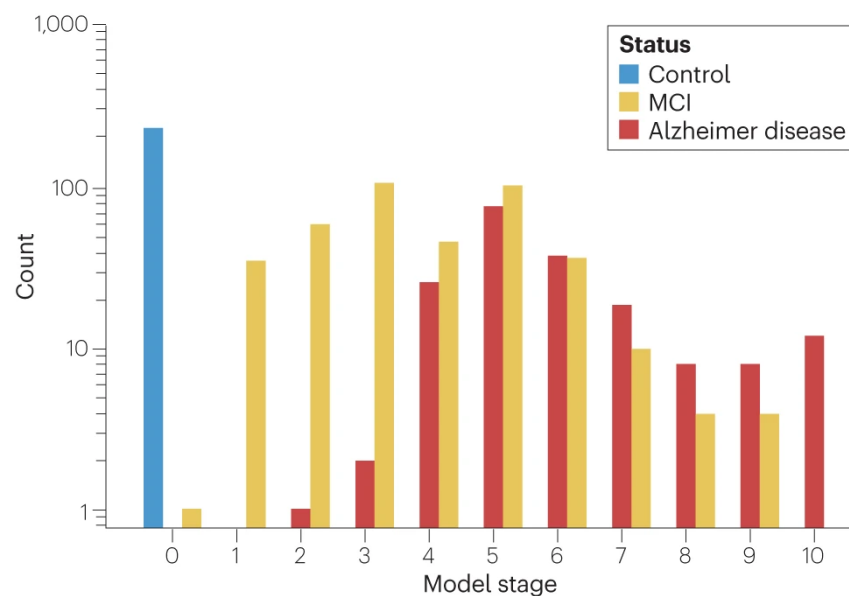
Vogel et al. Nat Med 2021

Applications

Data-driven modelling of neurodegenerative disease progression: thinking outside the black box

Alexandra L. Young^{1,2,6}, Neil P. Oxtoby^{1,6}, Sara Garbarino³, Nick C. Fox⁴, Frederik Barkhof^{1,5}, Jonathan M. Schott⁴ & Daniel C. Alexander¹

C Temporal stratification



Oxtoby et al. Frontiers AI 2022

Recap

Data-Driven Disease Progression Modelling

- Goes beyond “black box” approaches: “human insight + ML”
- Aids disease understanding at multiple scales
- Can support clinical decision making

Interested in Data-Driven Disease Progression Modelling?

<https://disease-progression-modelling.github.io>

Contributors



Igor Koval

Post-doc @ Paris Brain
Institute & Inria



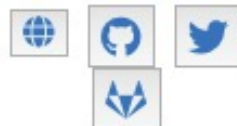
Neil Oxtoby

Senior Research Fellow @
UCL, UK



Sara Garbarino

Postdoctoral research
fellow @ UNIGE, Genova



Marco Lorenzi

Researcher @Inria Sophia
Antipolis, Université Côte
d'Azur, France



**Vikram
Venkatraghavan**

Postdoctoral researcher
@ Alzheimer Center,
Amsterdam UMC



Interested in Data-Driven Disease Progression Modelling?

<https://disease-progression-modelling.github.io>

Disease Progression Modelling

Search this book...

Disease Progression Modelling

MODELS

Introduction to DPM

Event Based Model

Disease Course Mapping

GP Progression Model

NOTEBOOKS

Overview

Disease Course Sequencing with the Event Based Model

Disease Course Mapping with Leaspy

GP Progression Model

CONFERENCES

Conferences

MICCAI 2021

ISBI 2021



Disease Progression Modelling



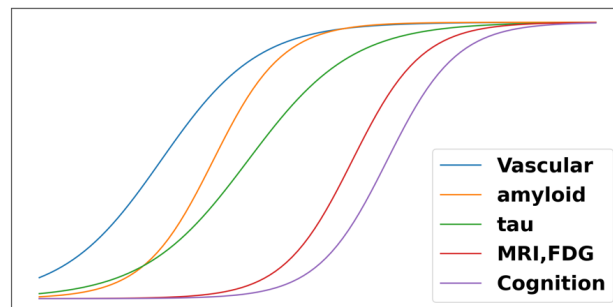
Contents

Disease Progression Modelling in a nutshell

Site content

The **Disease Progression Modelling** community unites medics with researchers and engineers across the physical and life sciences to tackle some of the biggest challenges of 21st-century medicine by harnessing the power of mathematics, computer science, and data.

—This website aims to serve as a portal for Disease Progression Modelling



The Alzheimer's Disease Progression Of Longitudinal Evolution Challenge



Predictive modelling challenge for Alzheimer's disease

tadpole.grand-challenge.org

TADPOLE SHARE: tadpole-share.github.io



EuroPOND



alzheimer's
association



Marinescu et al. [arxiv:1805.03909](https://arxiv.org/abs/1805.03909)

[MELBA Vol 1, 2021:19](#)

Acknowledgements



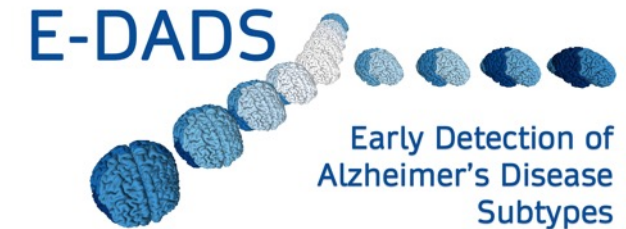
- UCL POND ucl-pond.github.io
Prof. Danny Alexander, Alexandra Young, et al.



- EuroPOND europond.eu



- E-DADS e-dads.github.io



- Collaborators, Data providers, Volunteers (patients & families)



Thank you

Data-driven Disease Progression Modelling: thinking outside the black box

Neil P. Oxtoby

Principal Research Fellow (research track Assoc. Prof.)
UCL Centre for Medical Image Computing



Data Science Lead, DEMON Network
demon dementia.com

