



EMORY

NELL HODGSON
WOODRUFF
SCHOOL OF
NURSING

Data-driven disease progression modeling

An application to Alzheimer's Disease

Alex Fedorov, Ph.D.
Research Assistant Professor
Nell Hodgson Woodruff School of Nursing
Emory University

January 23, 2024
Center for Data Science
All Hands Meeting



EMORY

NELL HODGSON
WOODRUFF
SCHOOL OF
NURSING

Data-driven disease progression modeling

An application to Alzheimer's Disease

Alex Fedorov, Ph.D.

Research Assistant Professor

Nell Hodgson Woodruff School of Nursing

Emory University

January 18, 2024

Reading group

Center for Translational Research
in Neuroimaging and Data Science (TReNDS)

Significance / Why?

Problem

Alzheimer's Disease

Alzheimer's Disease

Amyloid beta plaques
Tau neurofibrillary tangles
Atrophy
Cognitive decline

Patient History
Social Determinants

APOE4
PSIN1
PSEN 2
PSEN 2
APP
Down syndrome

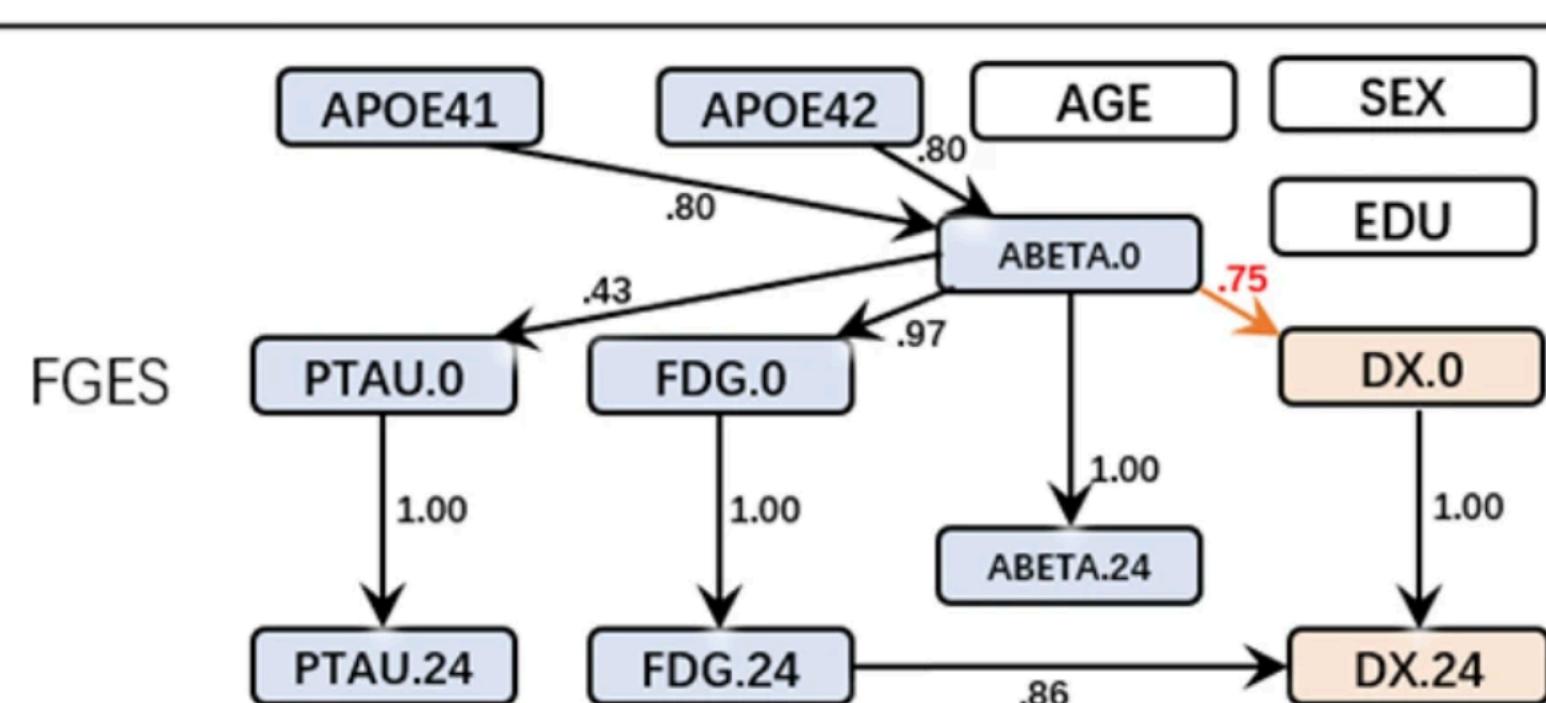
Cerebrospinal fluid (CSF)
Blood / Plasma Test

Structural MRI
Functional MRI
Diffusion MRI
FDG-PET
Amyloid Beta - PET
Tau - PET

Cognitive tests

Alzheimer's Disease

Amyloid beta plaques
Tau neurofibrillary tangles
Atrophy
Cognitive decline



Cerebrospinal fluid (CSF)

Blood / Plasma Test

Structural MRI

Functional MRI

Diffusion MRI

FDG-PET

Amyloid Beta - PET

Tau - PET

APOE4

PSIN1

PSEN 2

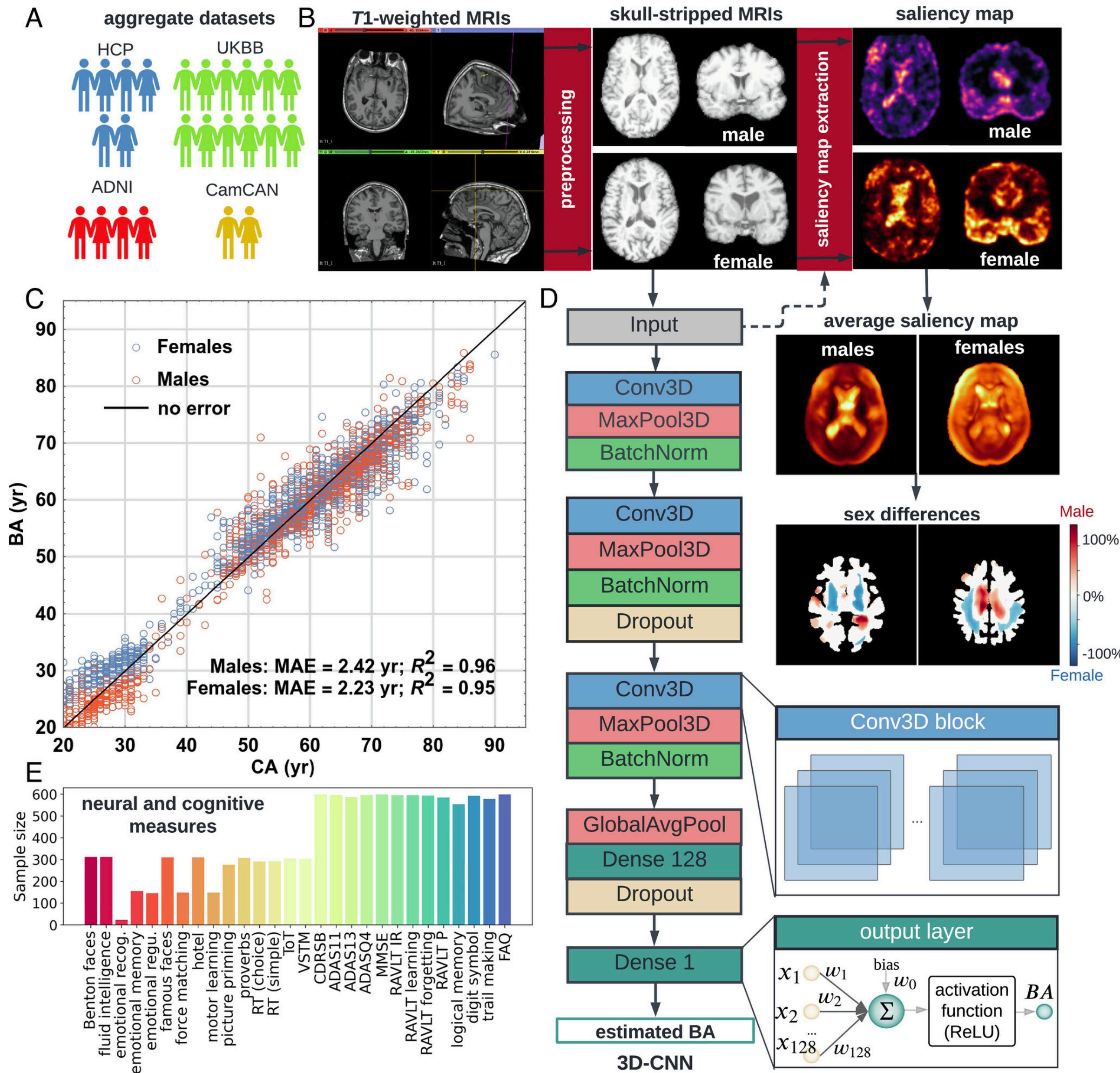
PSEN 2

APP

Down syndrome

Cognitive tests

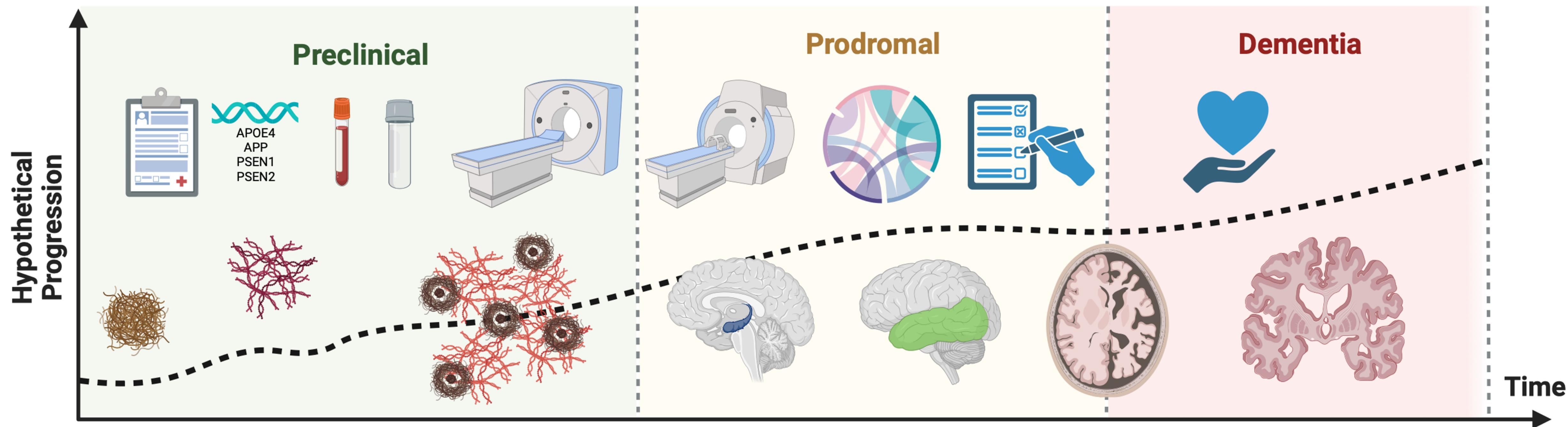
Typical Neuroimaging with Deep Learning



<https://www.pnas.org/doi/full/10.1073/pnas.2214634120>

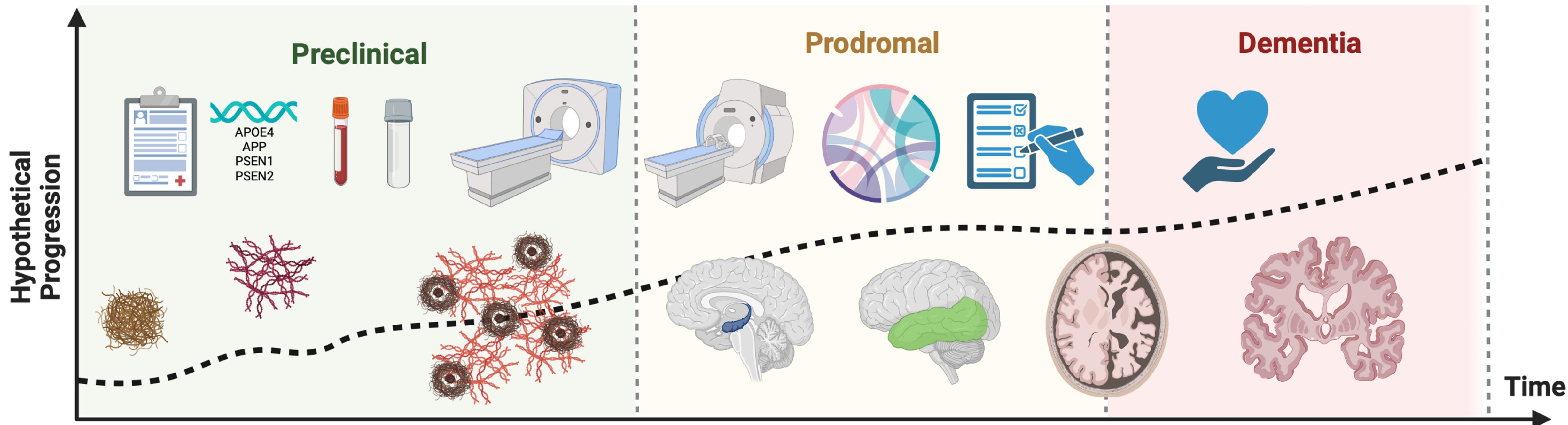
Hypothetical Progression

Alzheimer's Disease



Hypothetical Progression

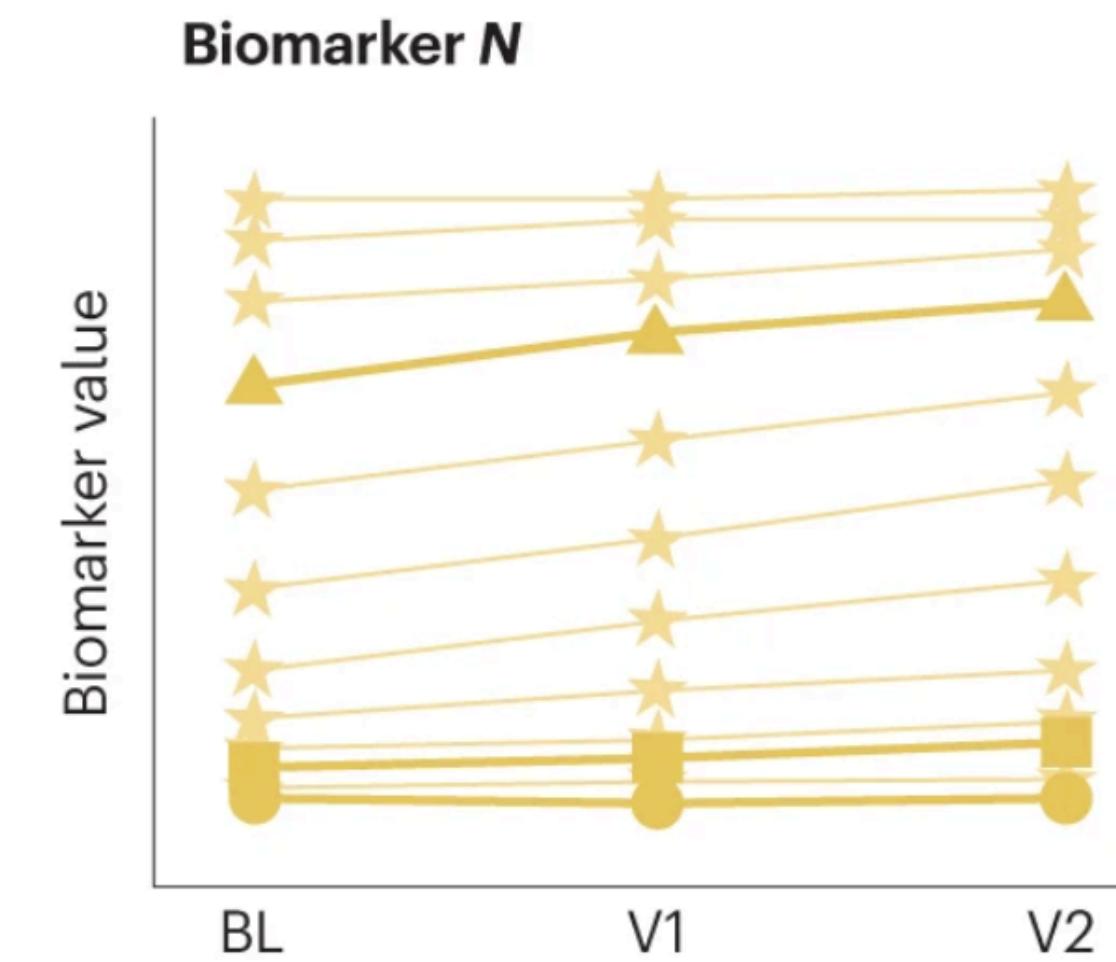
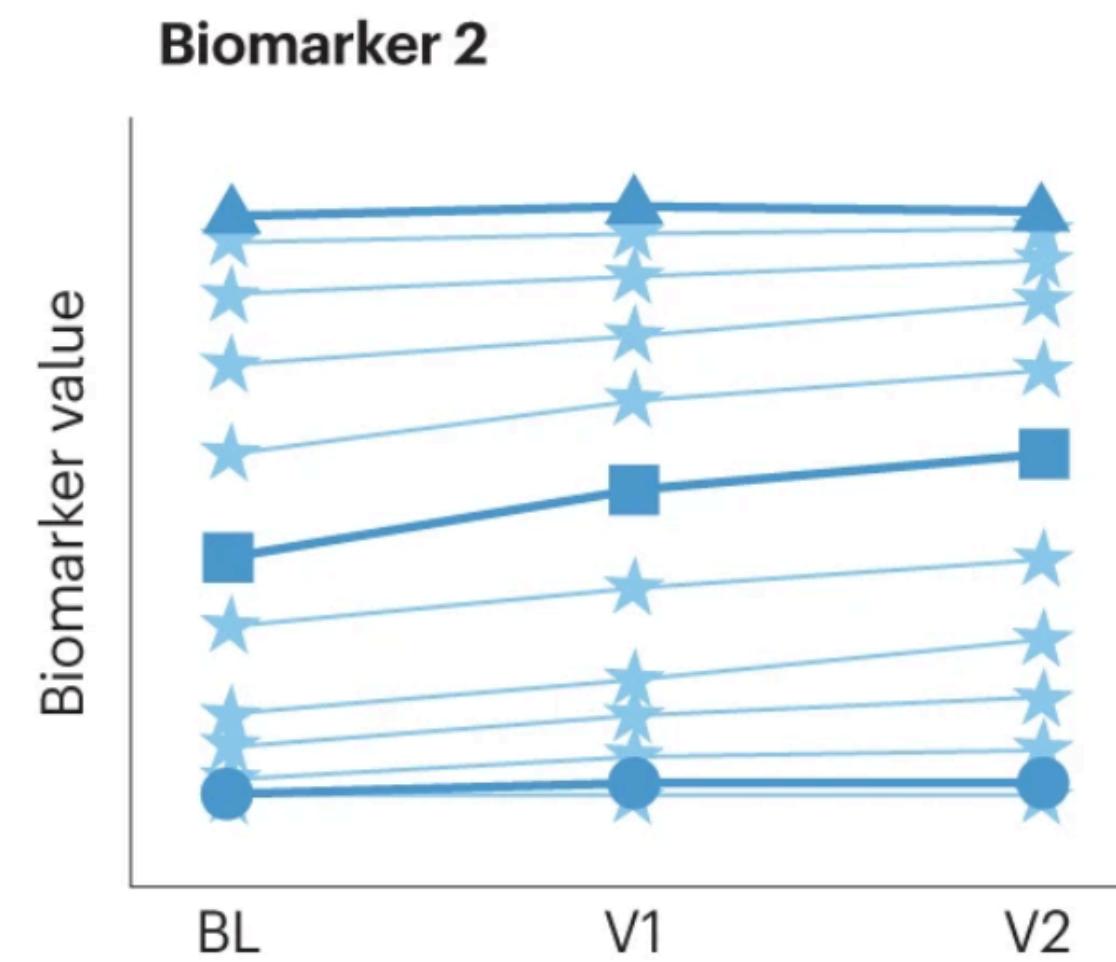
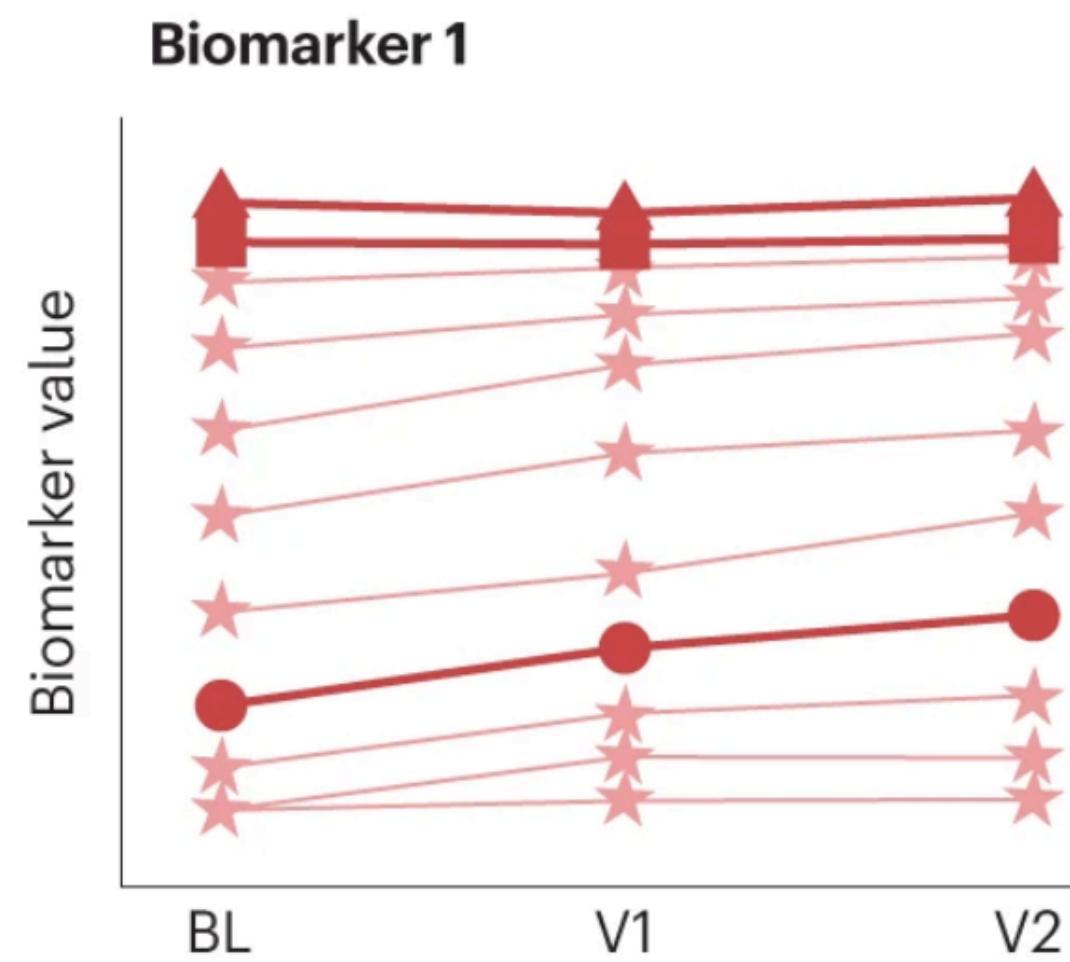
Alzheimer's Disease



- Disease is not static!
- Different diagnostic tools needed -> How to use them and what to prescribe?
- We want to treat early and monitor treatment

Data in longitudinal studies

Disease biomarker data indexed by visit



Legend:

- Individual 1
- Individual 2
- ...
△ Individual M

- Data in reality is
 - short-term longitudinal (usually only cross-sectional)
 - irregularly-sampled (sparser towards beginning)
 - missing observations
 - different order of observations procedures
 - biological age vs chronological age

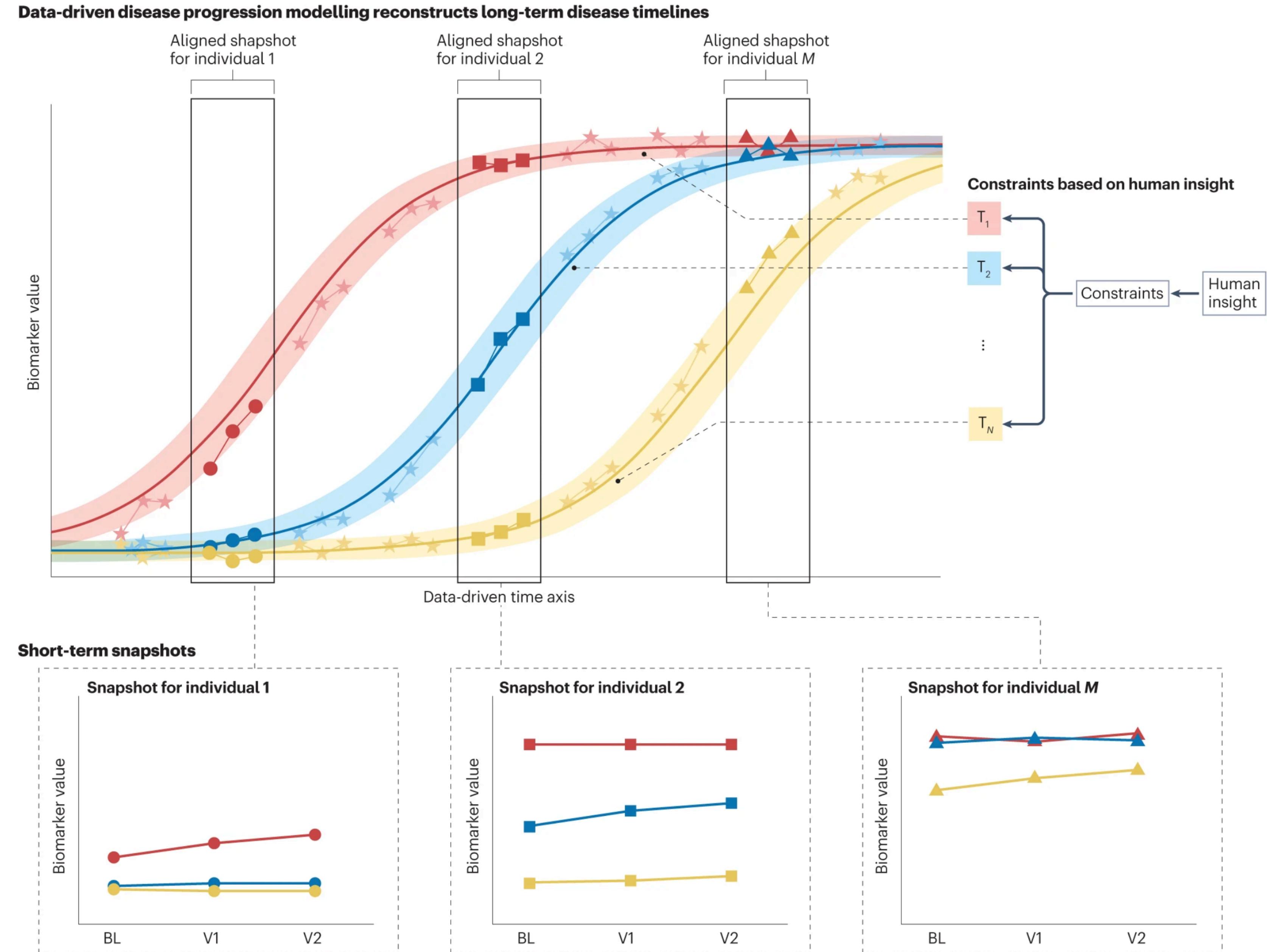
Data Collection

Challenges

- Decade-long assessment at scale
- Examinations can be: Inconvenient, Invasive, Expensive
- Earlier technology can be outdated
- Data Consistency
- Disease can be sporadic due to difficulty to identify before symptoms
- The data is under sampled during the early stages
- Heterogeneity of the disease

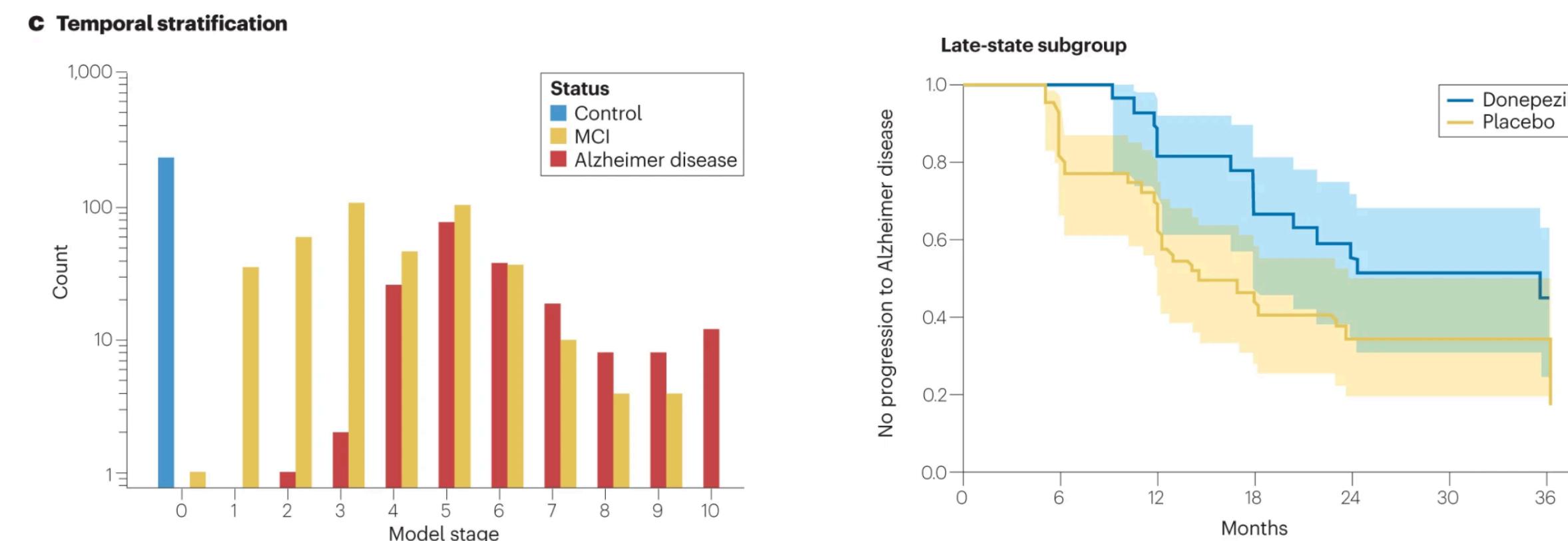
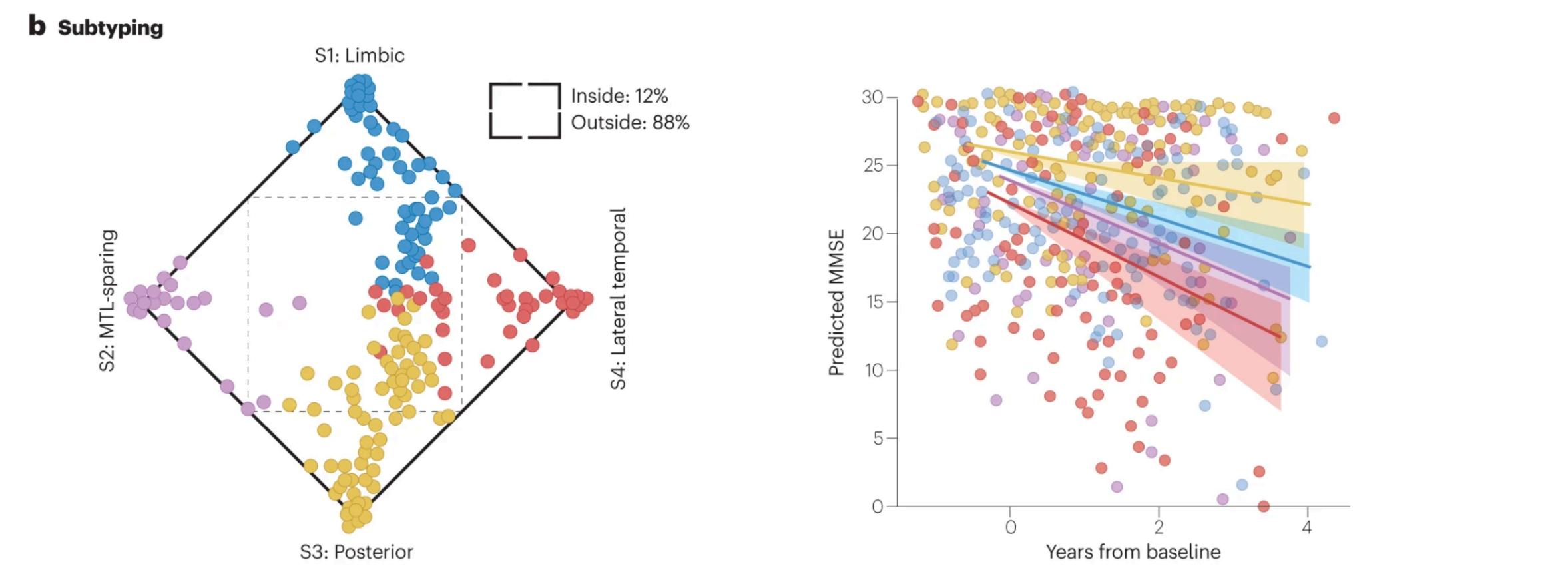
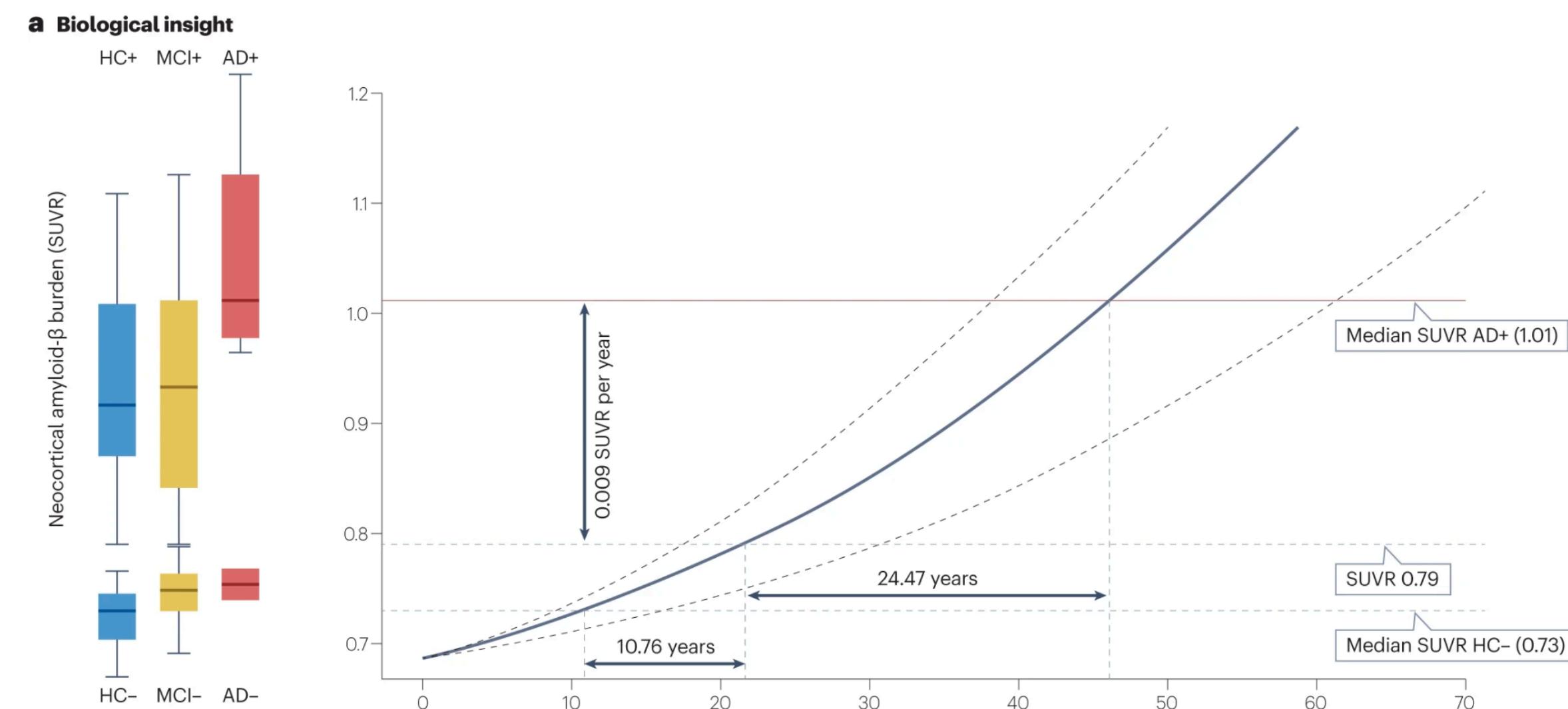
Disease Trajectories

- Data-driven disease progression models (**DDPM**)
- **Generative** model of how disease evolves over time
- Indexed by a **data-driven disease time axis**
 - **Expected average disease timeline**
 - Informed by **observations**



Applications

- Biological insight
- Patient staging
- Patient stratification
- Prognosis (not shown)



DDPM Taxonomy

Phenomenological

Modeling the trajectories without understanding underlying mechanism

Pathophysiological

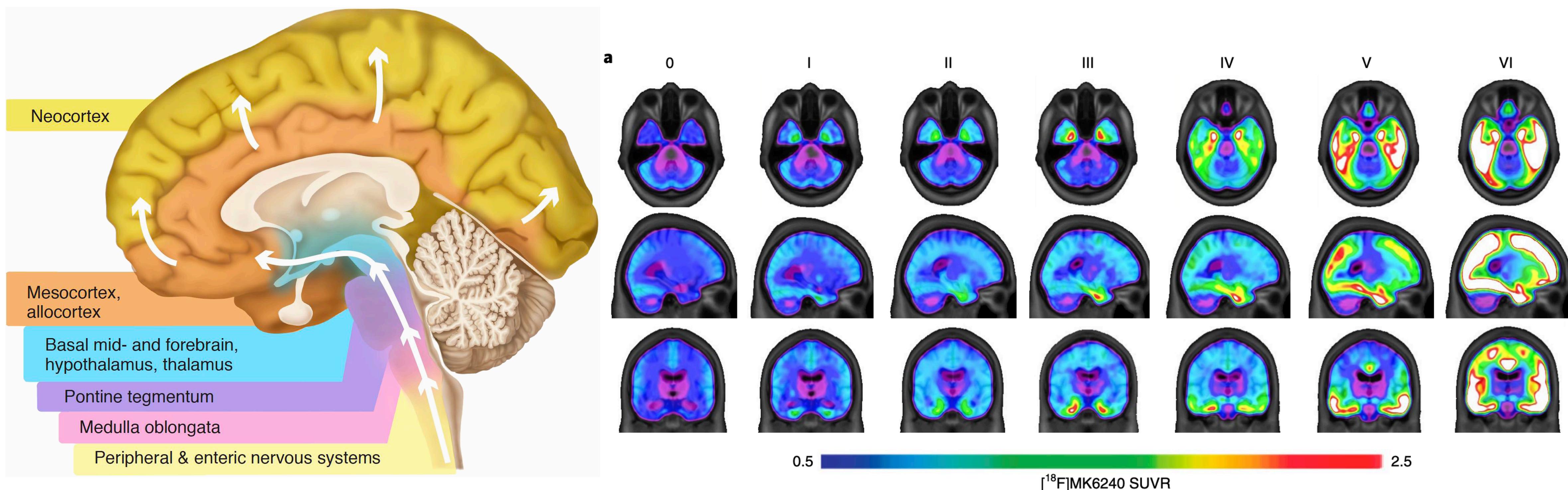
Explaining disease timelines in form of biological and physical processes (mechanisms)

Model	Utility	Input	Output	Constraint level
Phenomenological models				
Discrete	Data-driven disease patterns; temporal stratification; progression prognosis	Scalar (for example, spreadsheet)	Sequence of states	Low (for example, monotonic)
Continuous	Data-driven disease patterns; temporal stratification; progression prognosis	Scalar (for example, spreadsheet)	Trajectories	Low (for example, monotonic)
Spatiotemporal	Enhanced spatial localization of disease changes	Imaging data	Trajectories	Moderate; spatial constraints
Subtyping	Subtype stratification	Inherited from chosen phenomenological model	Inherited from chosen phenomenological model	Inherited from chosen phenomenological model
Pathophysiological models				
Network	Evaluating competing mechanistic hypotheses	Brain maps of connectivity	Expected sequence of pathology progression for evaluation against imaging changes	Very high; connectivity entirely defines pathology pattern
Dynamical systems	Pathophysiological parameter estimates	Brain maps; imaging	Trajectories; pathophysiological parameters	Relatively high; estimate only a few key pathophysiological parameters
Mechanistic combinations	Combinations and, possibly, interactions	Inherited from chosen pathophysiological model	Inherited from chosen pathophysiological model	Moderate; varies according to the number of mechanisms and pathophysiological parameters for each mechanism

Braak staging

Neuropathological staging systems

- The degree of pathology in Parkinson's and Alzheimer's diseases



Neuropathological biomarker models

Dynamic biomarkers of the Alzheimer's pathological cascade

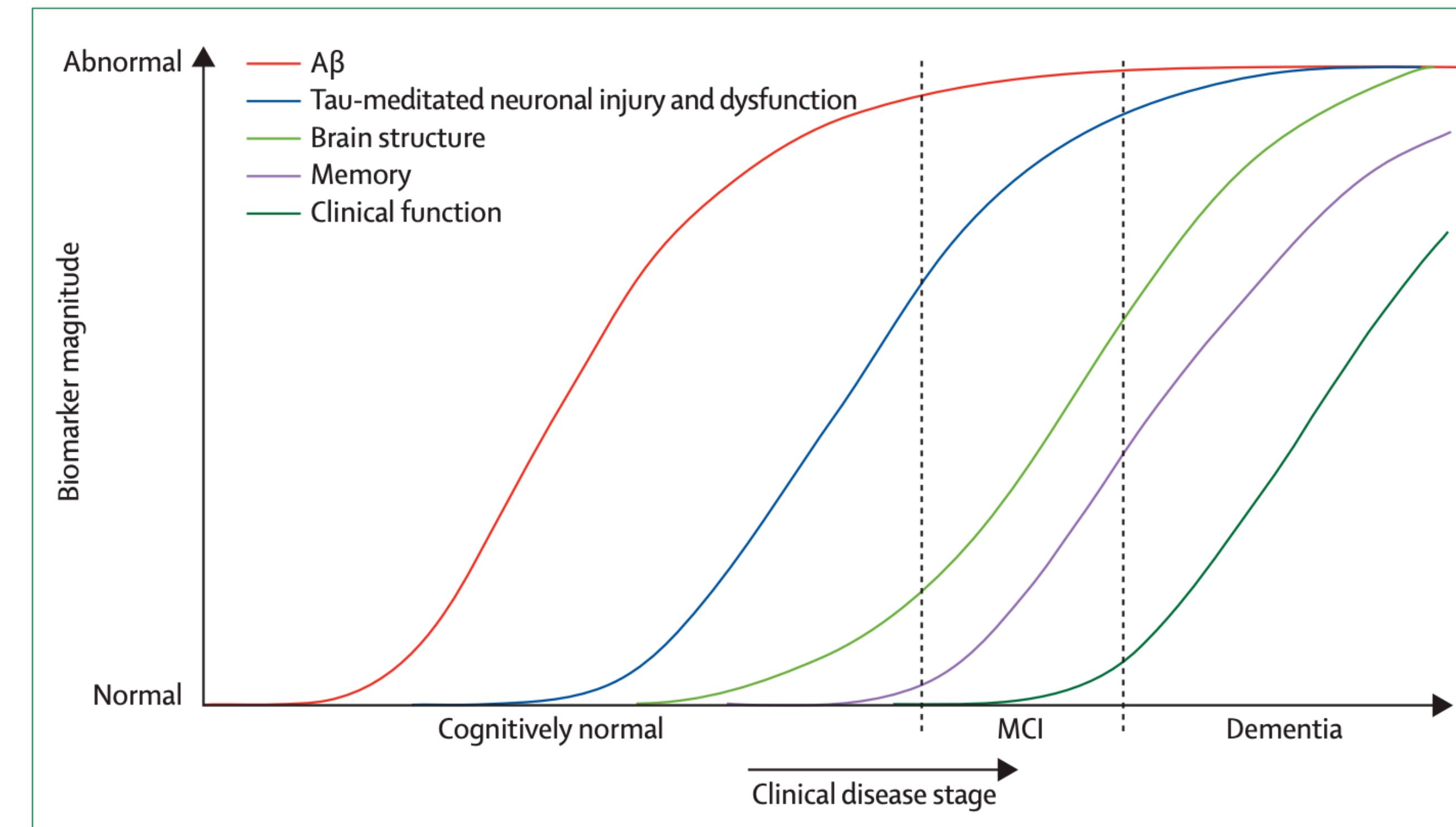
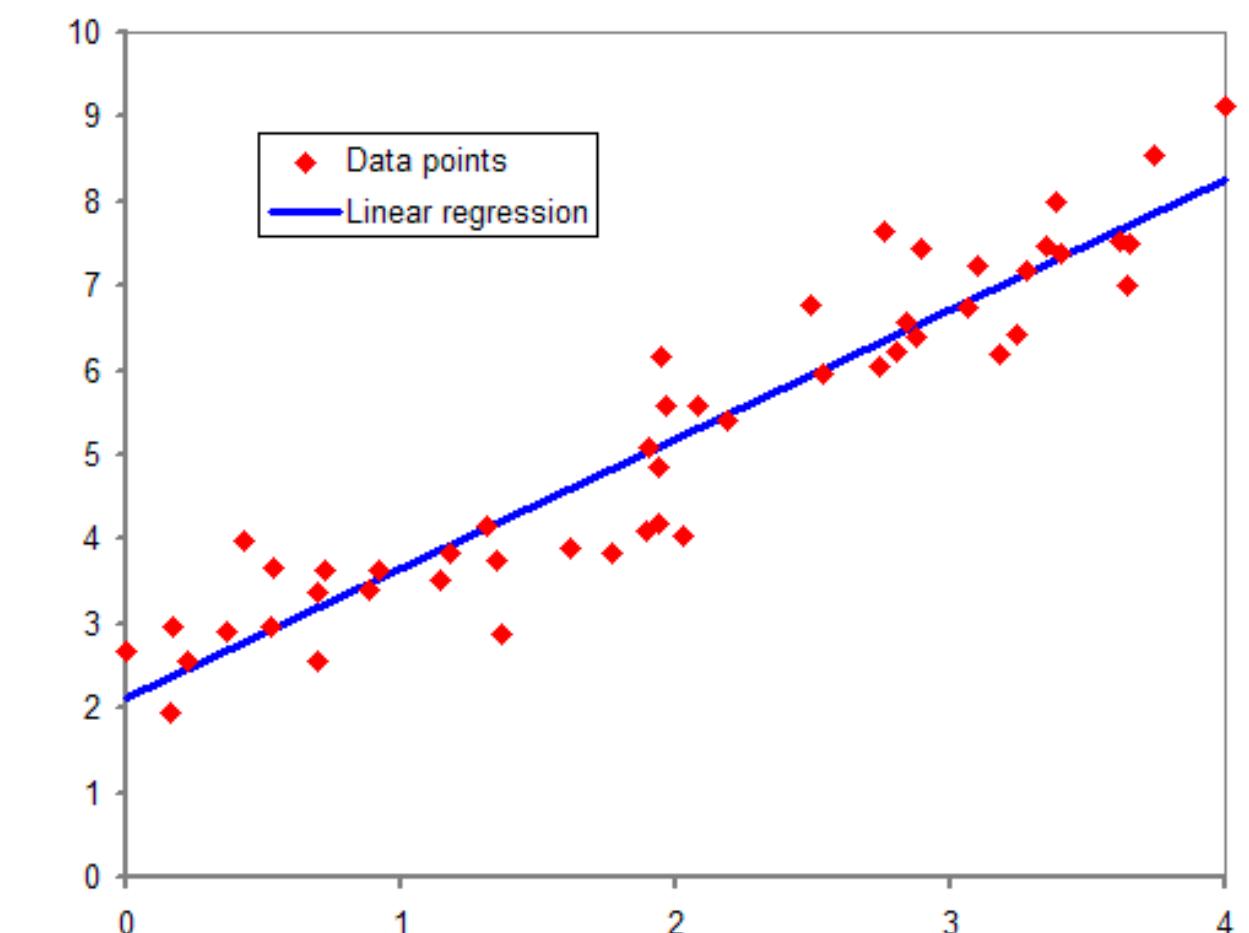


Figure 2: Dynamic biomarkers of the Alzheimer's pathological cascade

A β is identified by CSF A β_{42} or PET amyloid imaging. Tau-mediated neuronal injury and dysfunction is identified by CSF tau or fluorodeoxyglucose-PET. Brain structure is measured by use of structural MRI. A β = β -amyloid. MCI=mild cognitive impairment.

Regression is too simple

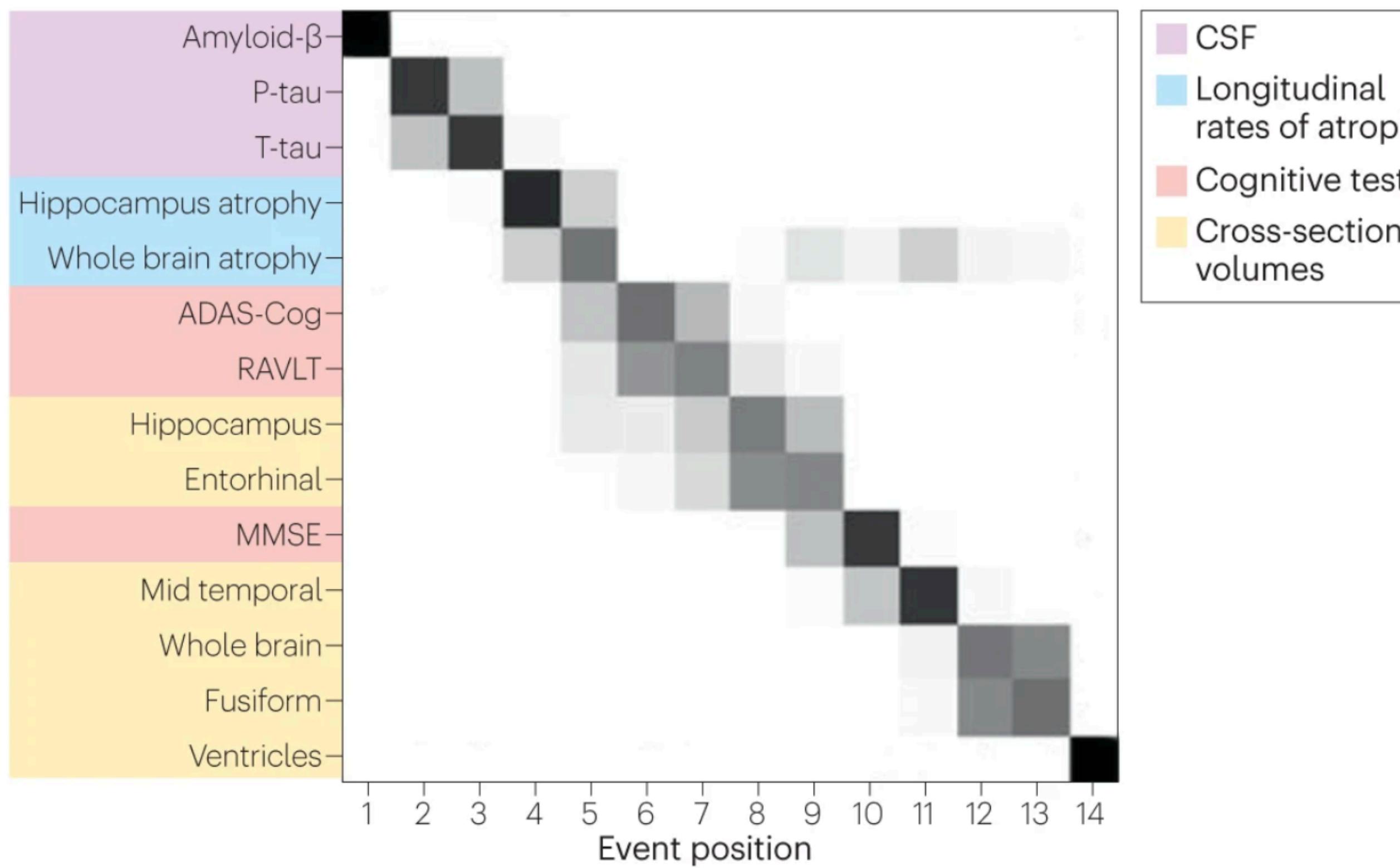
- Chronological age -> Age onset if variable -> Biological vs. Chronological age
- Clinical staging -> Crude -> Does not account well for spectra (heterogeneity) and might be useless for a particular stage
- Biomarker indexing -> Some biomarkers only sensitive to a particular stages
- Expected age of onset -> Genetic vs. Sporadic
- Time to conversion -> Early stage is usually under sampled



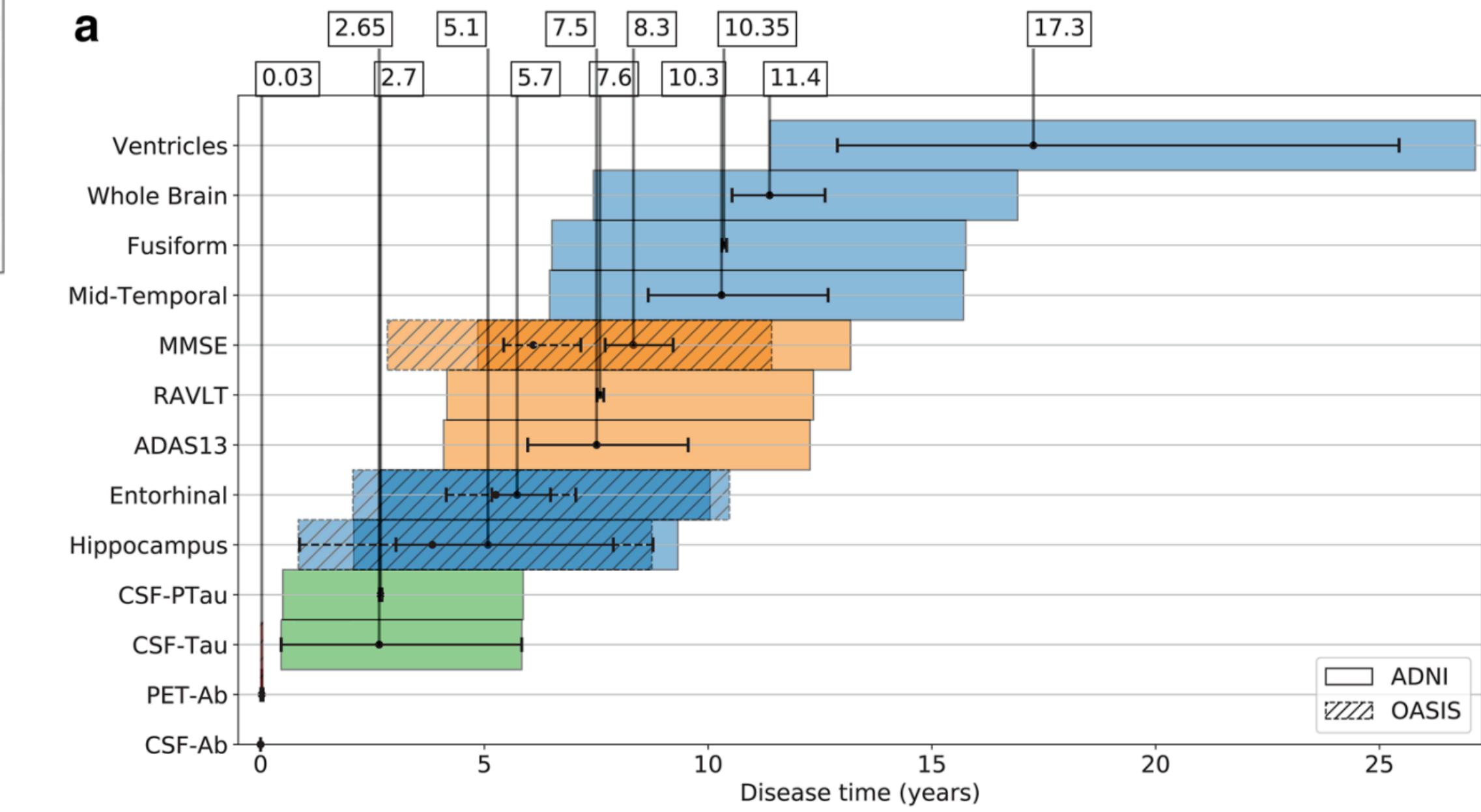
Event Based Model (EBM)

Phenomenological models

a



a

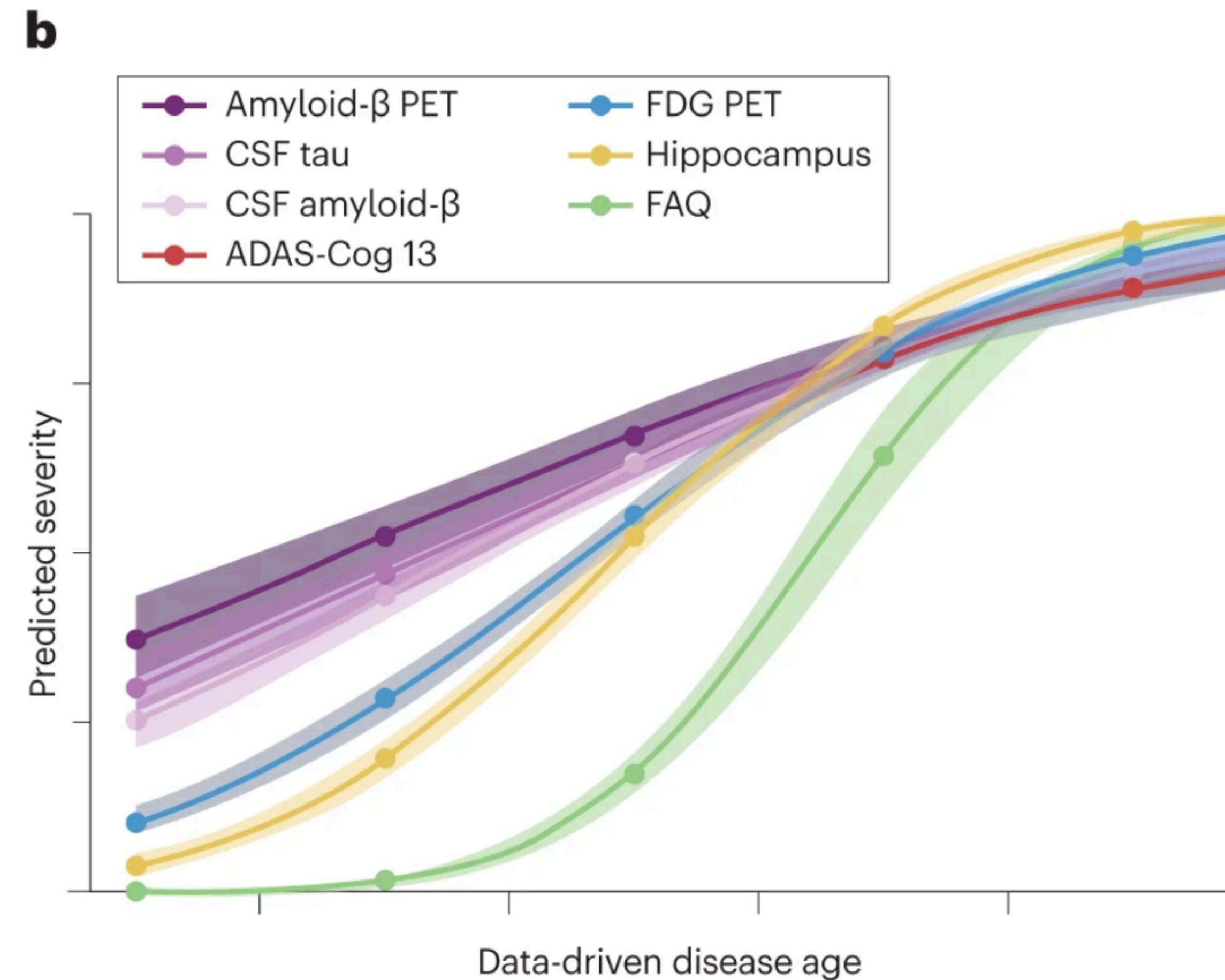


<https://academic.oup.com/brain/article/137/9/2564/2848155>

https://direct.mit.edu/imag/article/doi/10.1162/imag_a_00010/117183/The-temporal-event-based-model-Learning-event

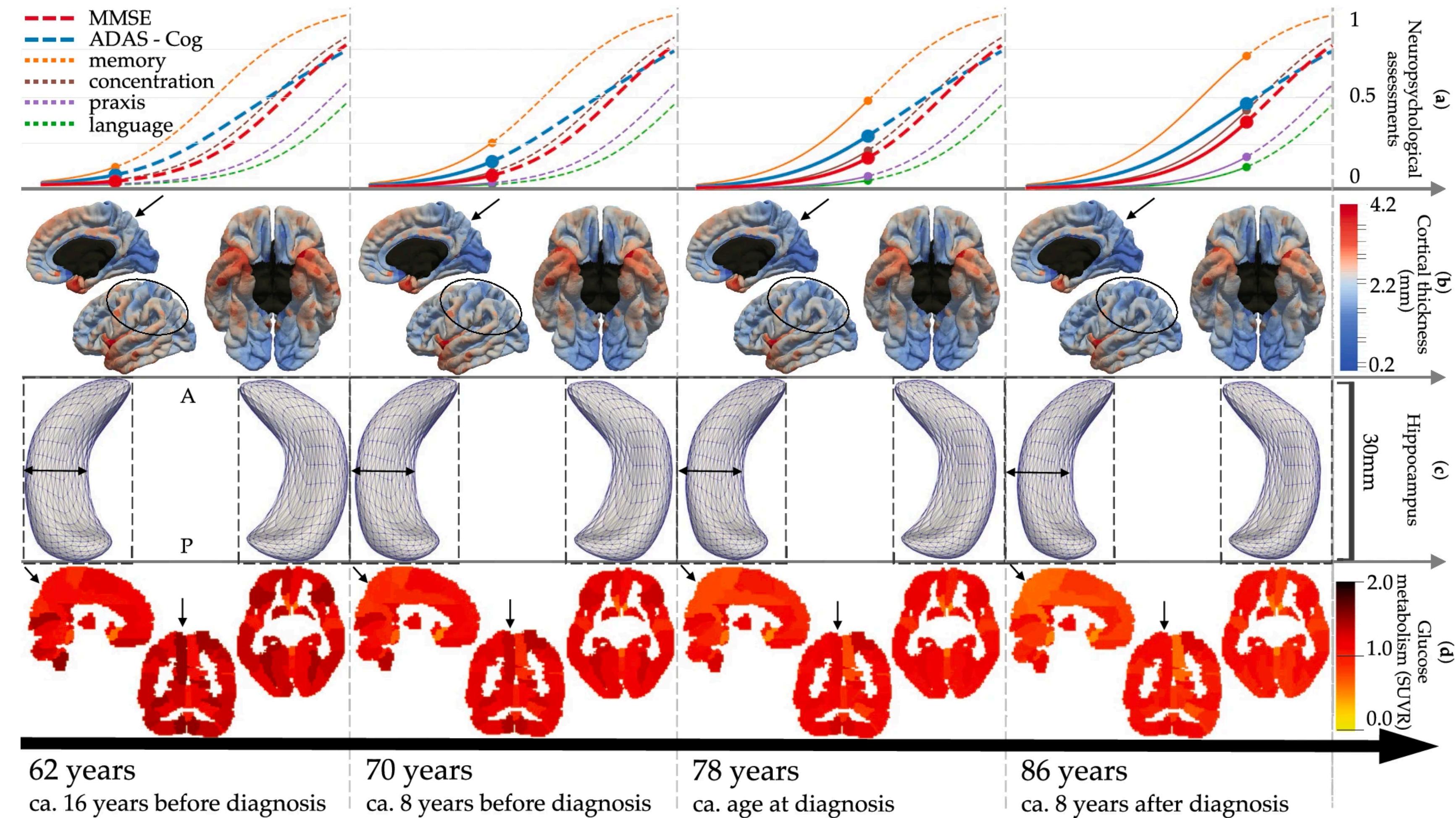
Latent time regression

Phenomenological models



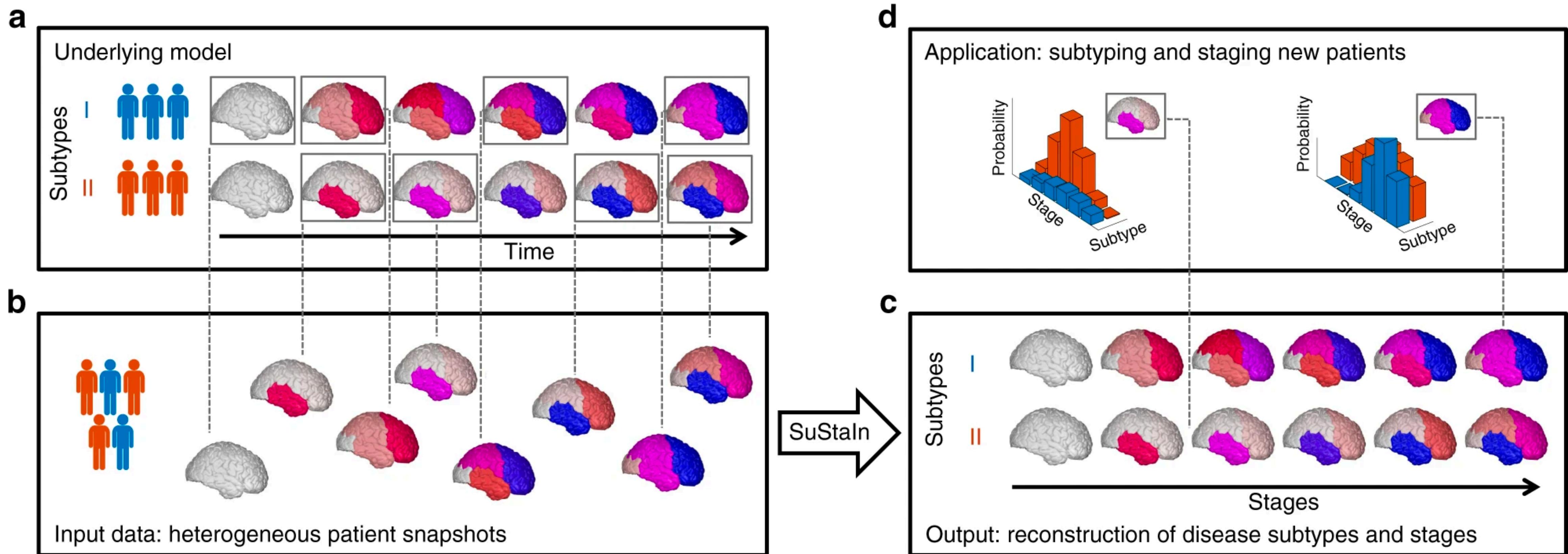
Spatiotemporal Models

Phenomenological models



Subtype and Stage Inference (SuStain)

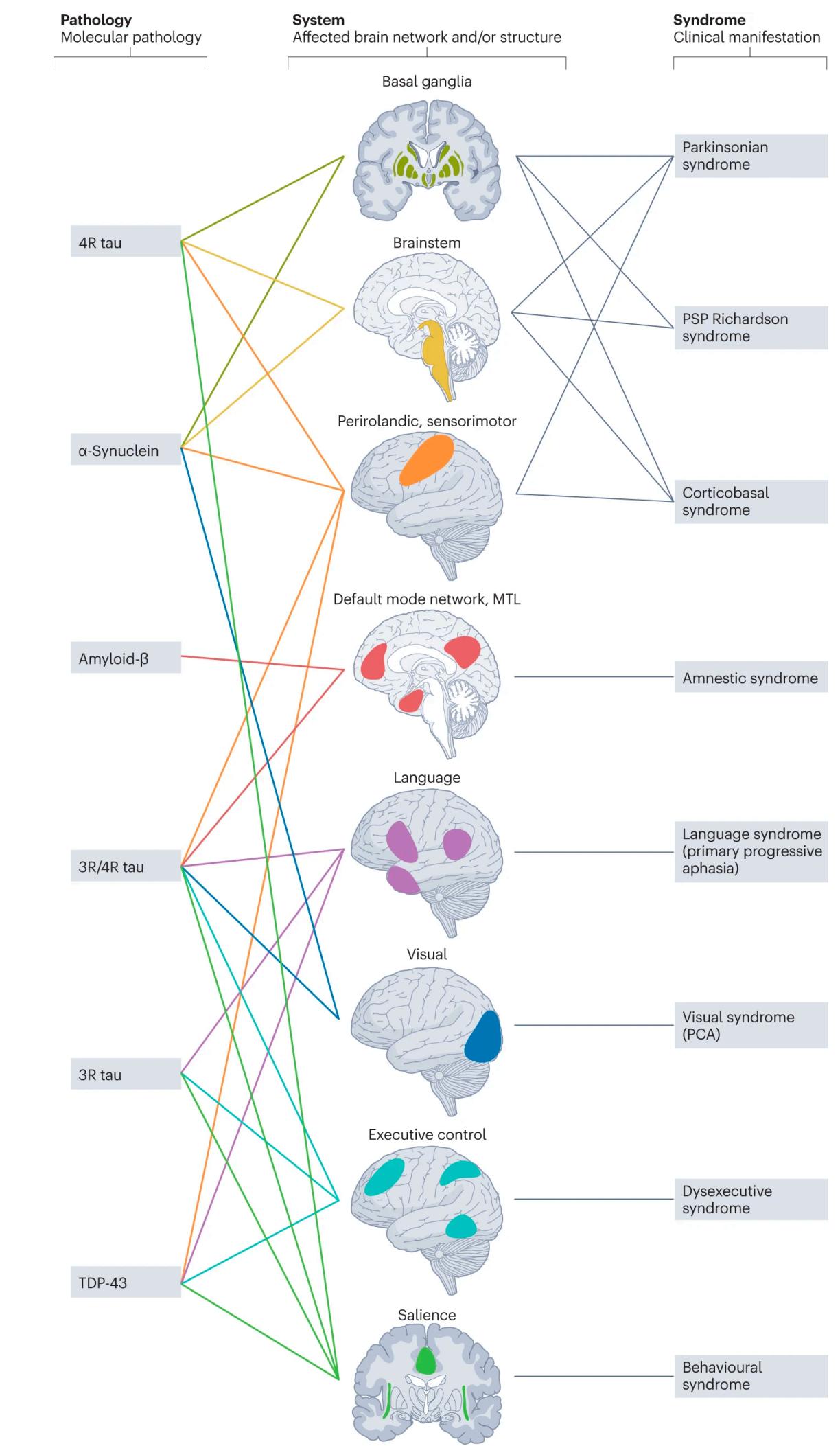
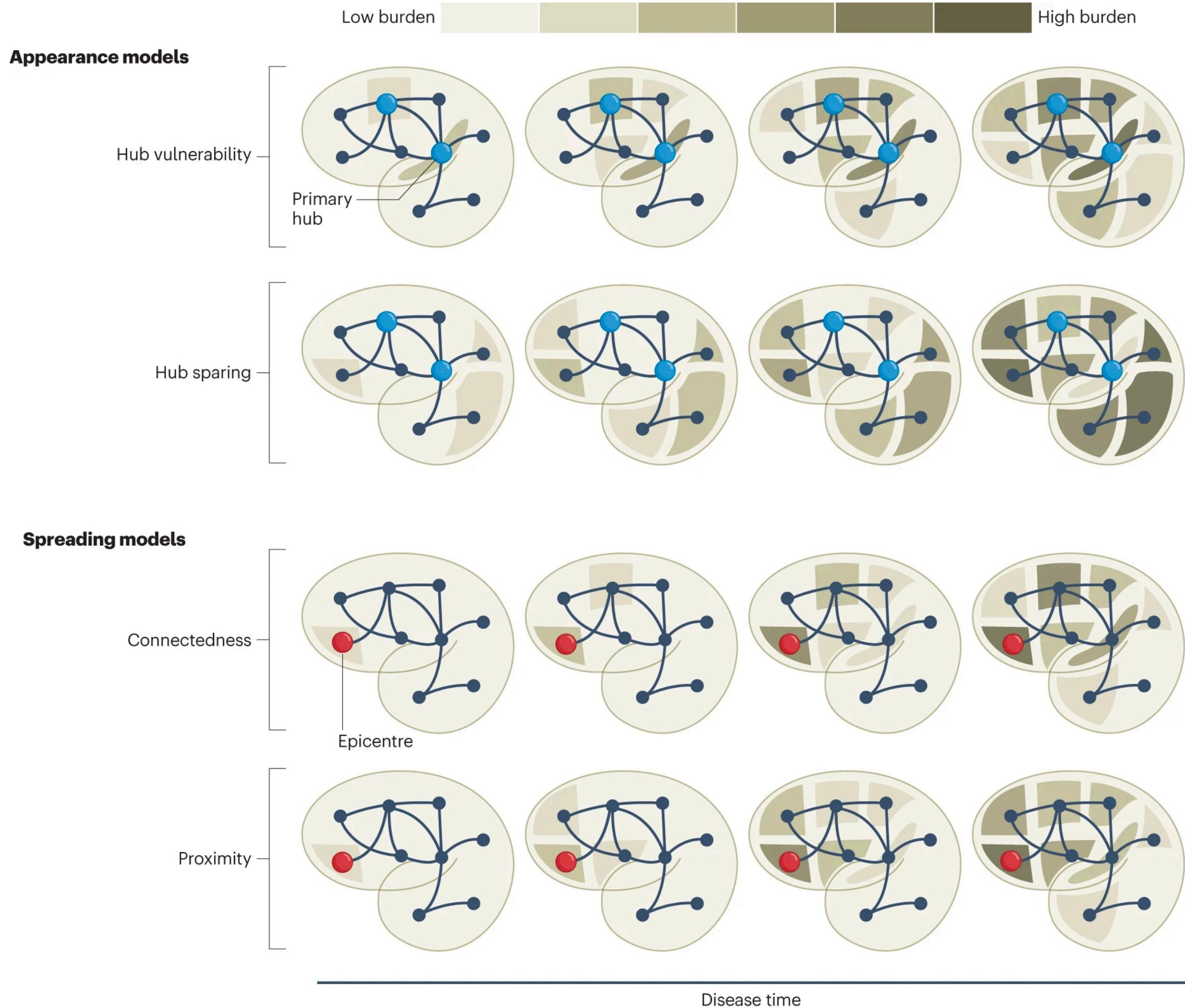
Phenomenological models



<https://www.nature.com/articles/s41467-018-05892-0>

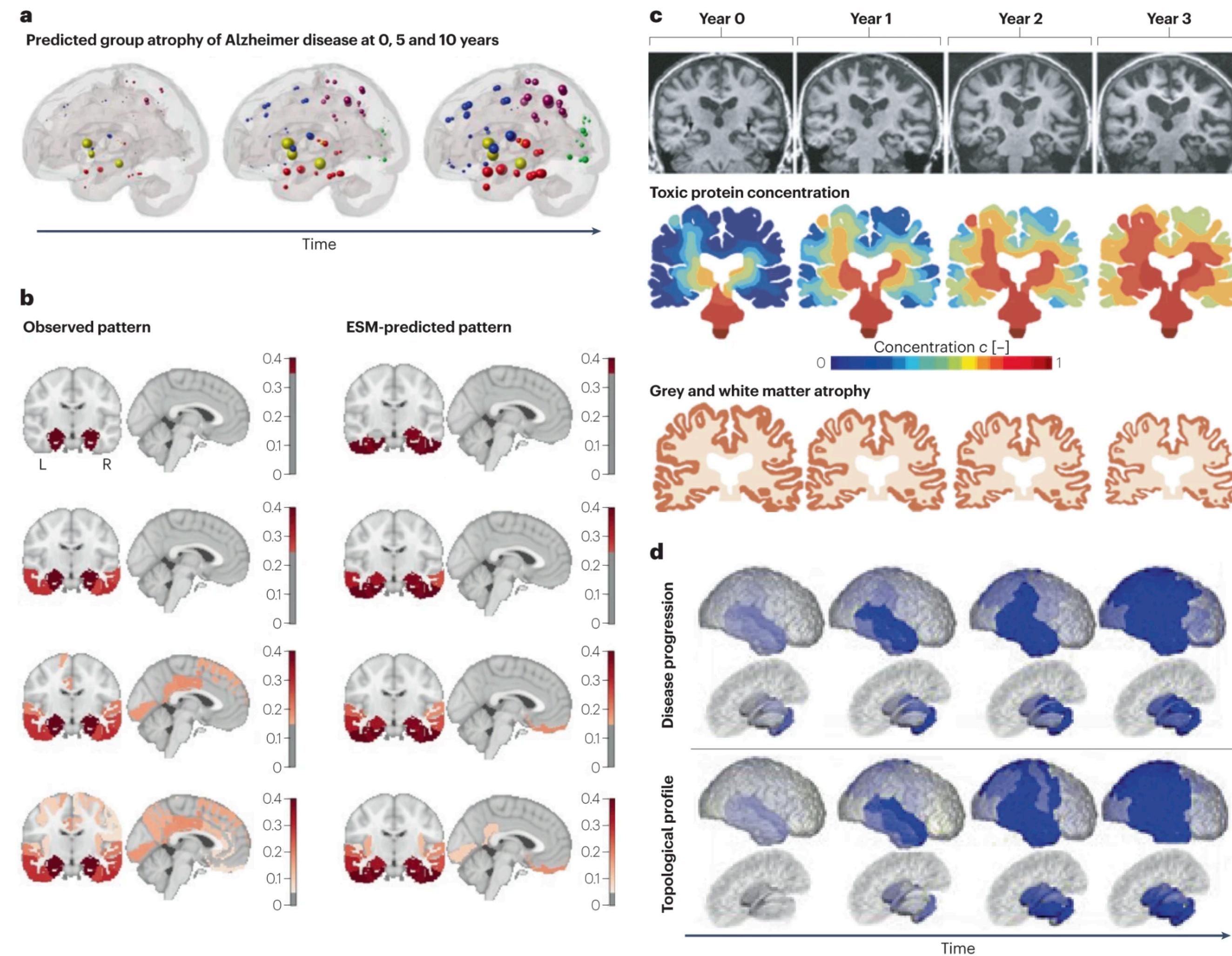
Appearance and Spreading Models

Pathophysiological models



Dynamical Systems Models

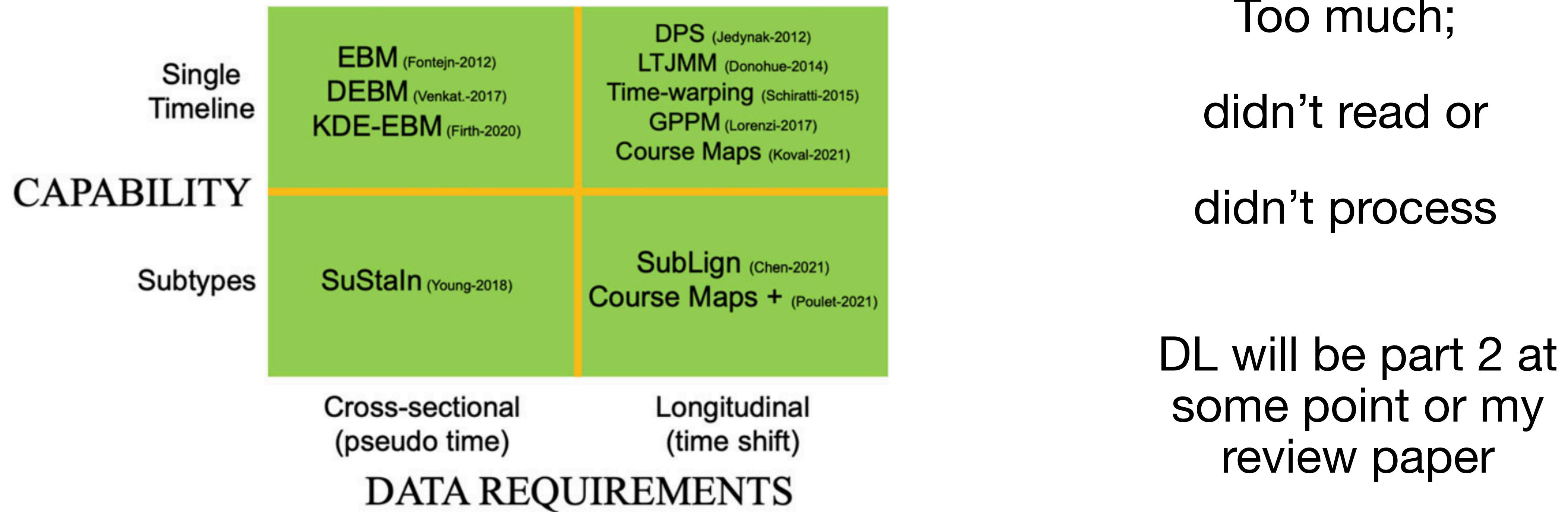
Pathophysiological models



Future technical Directions

- **Automatic Feature Extraction:** Multimodal, Interpretable
- **Treatment effects** (I do not understand / read this topic enough)
- **OMICS:** genetic risks, progression patterns and speed, early treatment before irreversible consequences
- **Integrative models across scales**
- **Age Effects:** Normal Aging (Biological vs Chronological) vs. Neurogenerative “Aging” (Disease -/-> Faster Aging)
- **Multimorbidity**
- **Mixed Pathology:** Multiple neurodegenerative diseases at the same time, or some biomarkers can be similar.
- **More mechanisms:** metabolic, cardiovascular, neuroinflammation, etc (Need MD )
- These models can be applied to **different domains:** psychiatric disorders, or any other organ disease
- **Spectra:** The nature of diseases can be very heterogeneous, hence we might oversimplify by using **distinct** groups.

There is much more in DDPM with ML and DL



https://www.ncbi.nlm.nih.gov/books/NBK597485/pdf/Bookshelf_NBK597485.pdf

Three ideas I am thinking...

Multimodal Representation Learning

Automatic fine-grained
feature extraction from
ANY modality

Progression Modelling

Can these
representation support
previous progression
insights as manual
features?

Decision Making

Can these
representations
support decision
making in care?

Research

Potential

Innovation / How?

Approach / What?