gsk.com

Bayesian life-course modelling of Alzheimer's Disease progression

Dave Lunn, Oana Petrof, Aris Perperoglou

PSI 2025 Conference, London, June 8-11

Acknowledgments

- The following work is proof-of-concept, inspired by…

- Raket, LL (2020) Statistical Disease Progression Modeling in Alzheimer Disease, *Frontiers in Big Data,* **3**

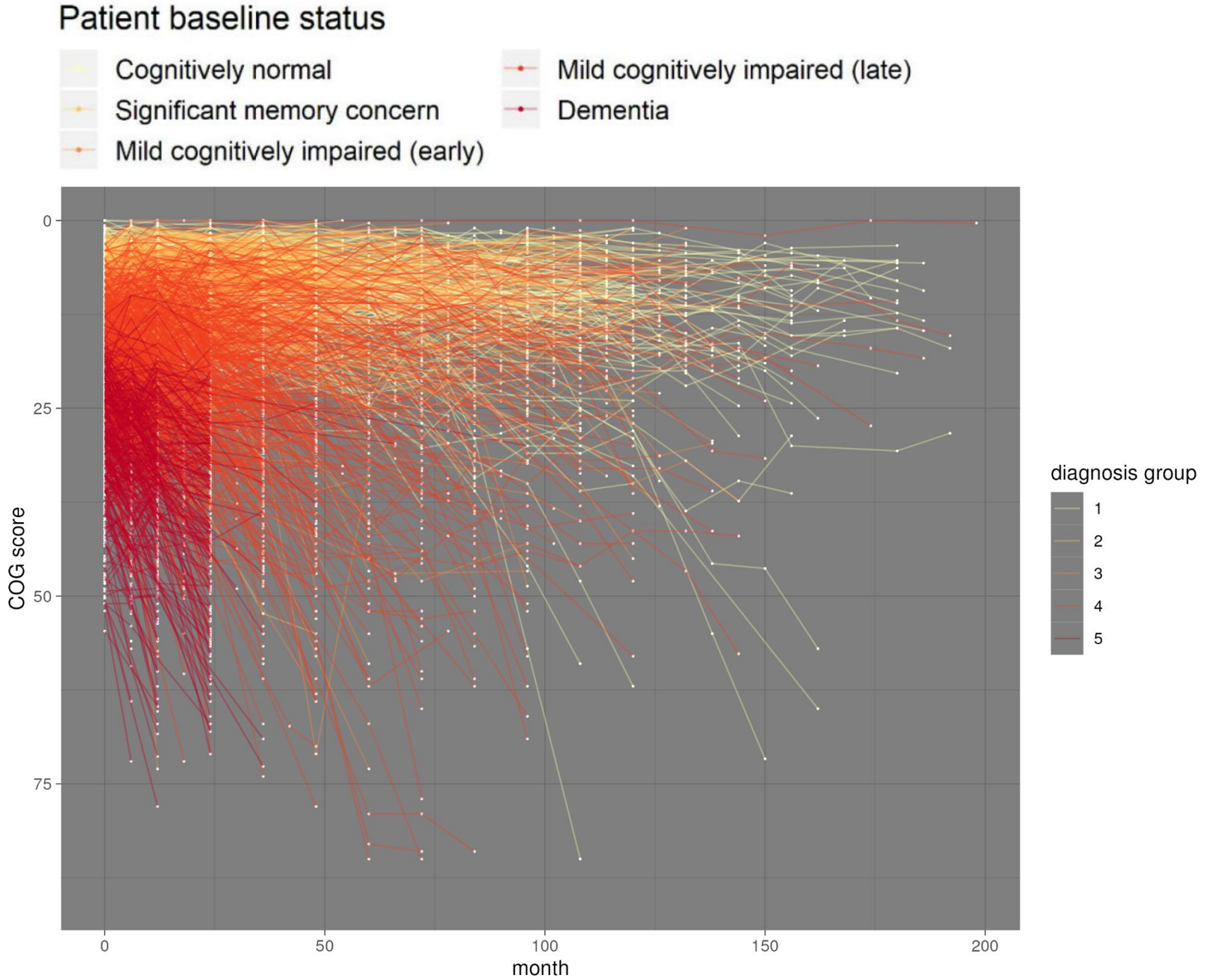
- Oana Petrof (SDS-IH) currently carrying the work forwards

ADNI data

- The Alzheimer’s Disease Neuroimaging Initiative (ADNI) is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of Alzheimer’s disease (AD)…
- We focus on ADAS-Cog 13 (85 point score) as measure of disease progression:
  - ~11,000 longitudinal observations
  - ~2400 individuals
  - Between 1 and 17 observations each
- Five baseline-diagnosis groups:
  1. Cognitively Normal
  2. Significant Memory Concern
  3. Early Mild Cognitive Impairment
  4. Late Mild Cognitive Impairment
  5. Alzheimer's Disease

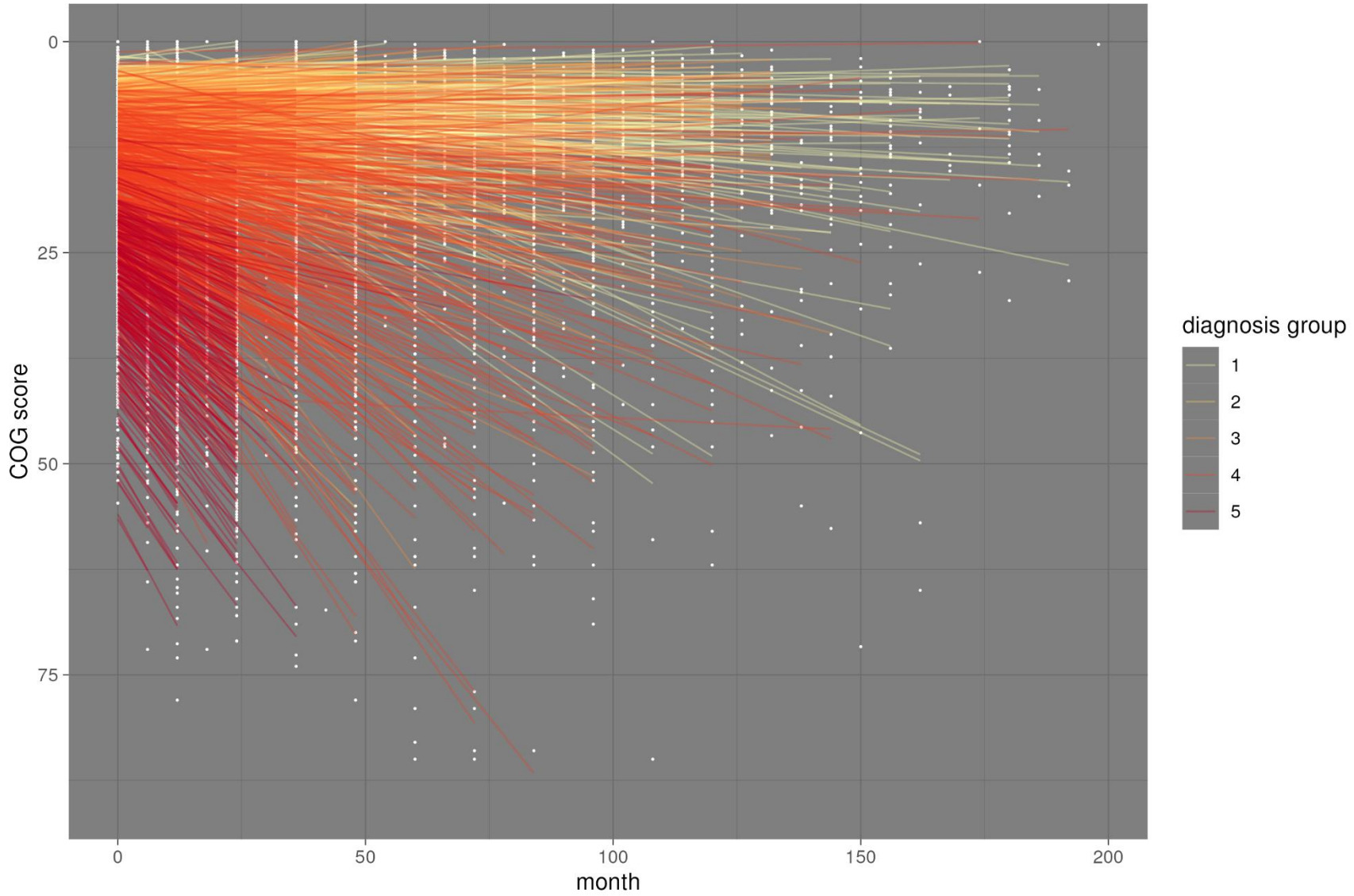
Raw data

- Larger scores and steeper slopes for more severe baseline diagnoses
- Some “curvature” apparent, so linear model may not be appropriate



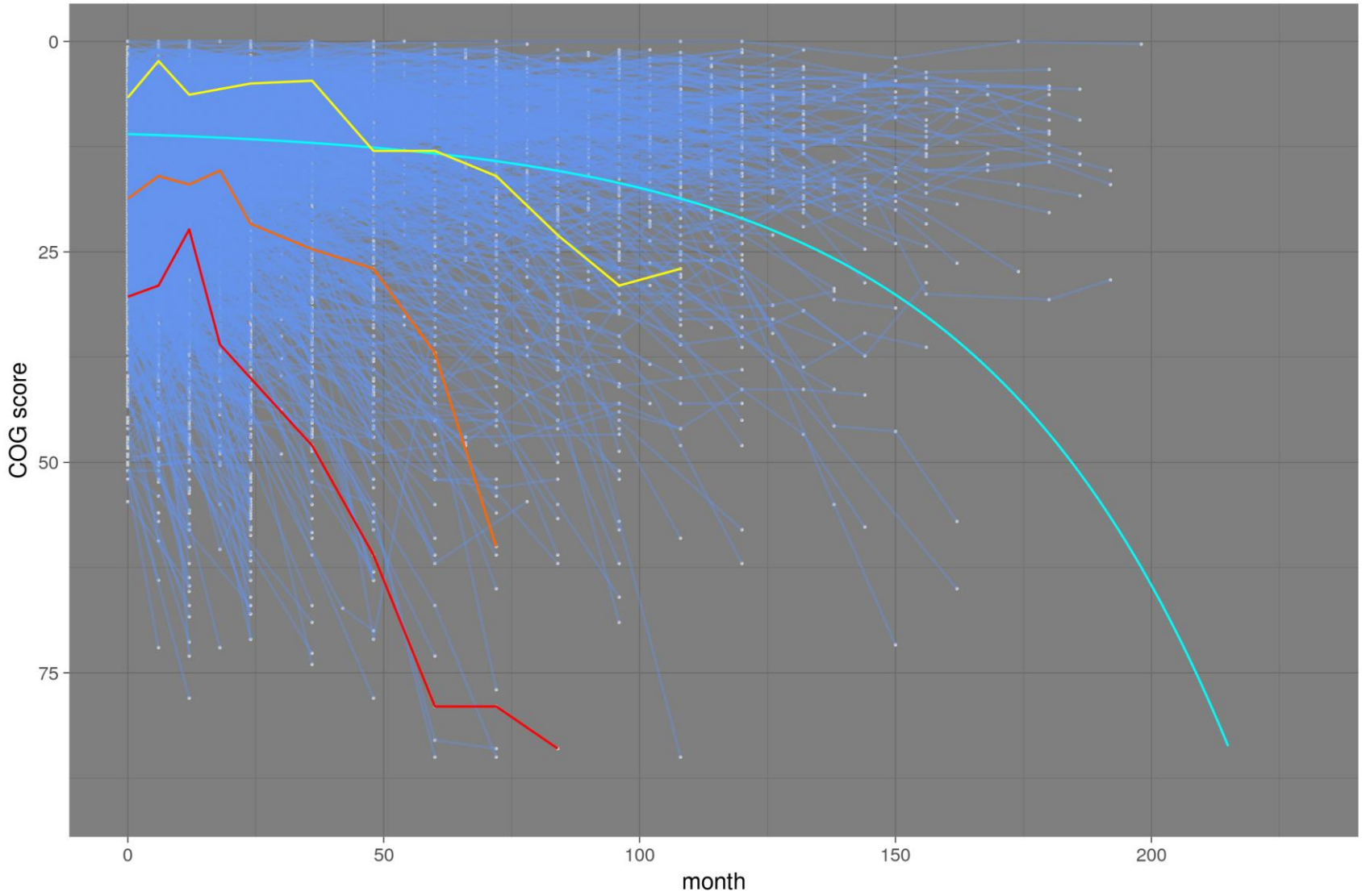
Linear Mixed Effects model

- Model assumes everyone following a different trajectory
- Perhaps better to assume everyone following a single, non-linear trajectory…
- … but they are observed at different points along that trajectory
- Model framed in terms of disease time rather than study time

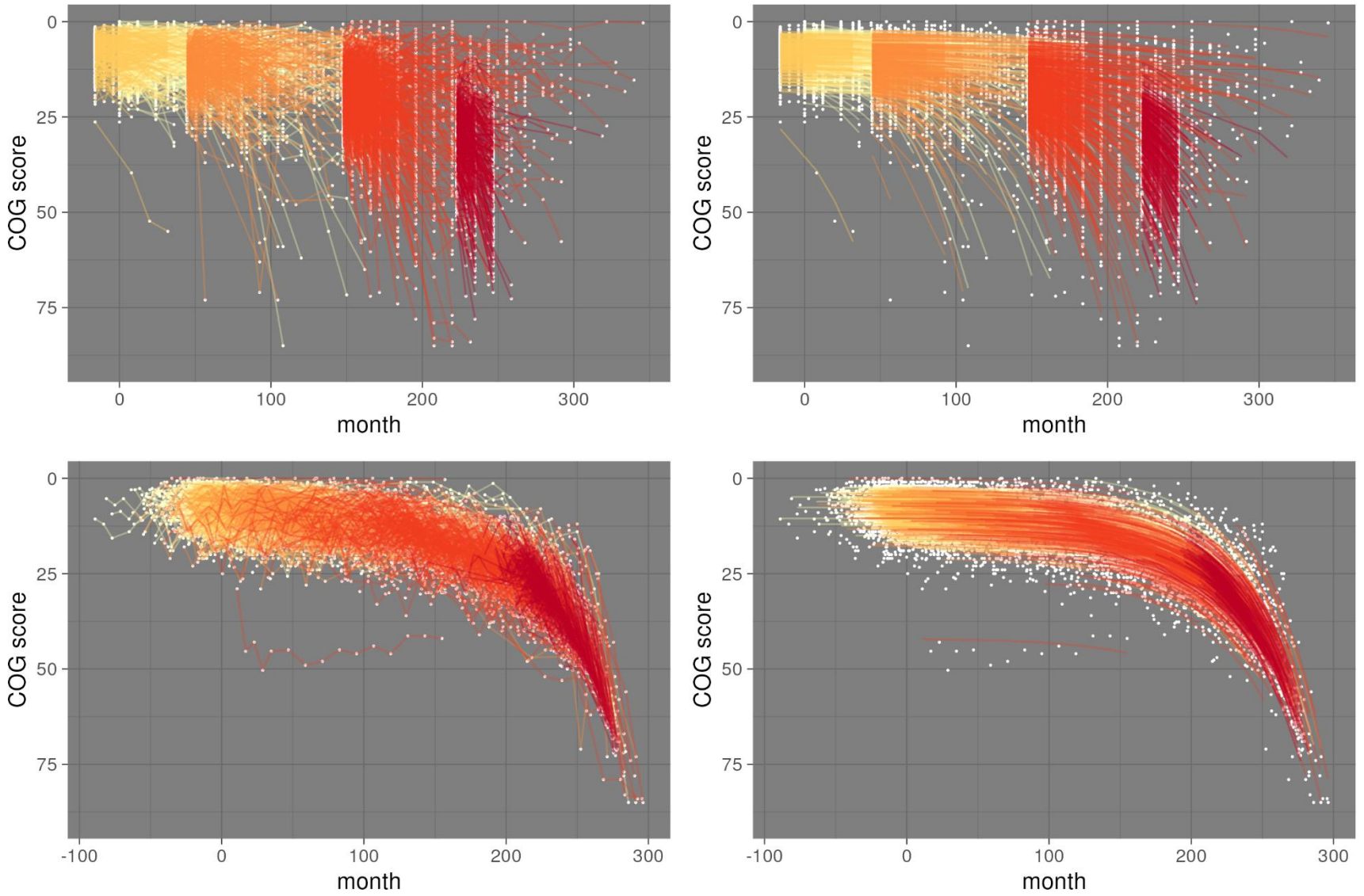


Disease-time (life-course) model

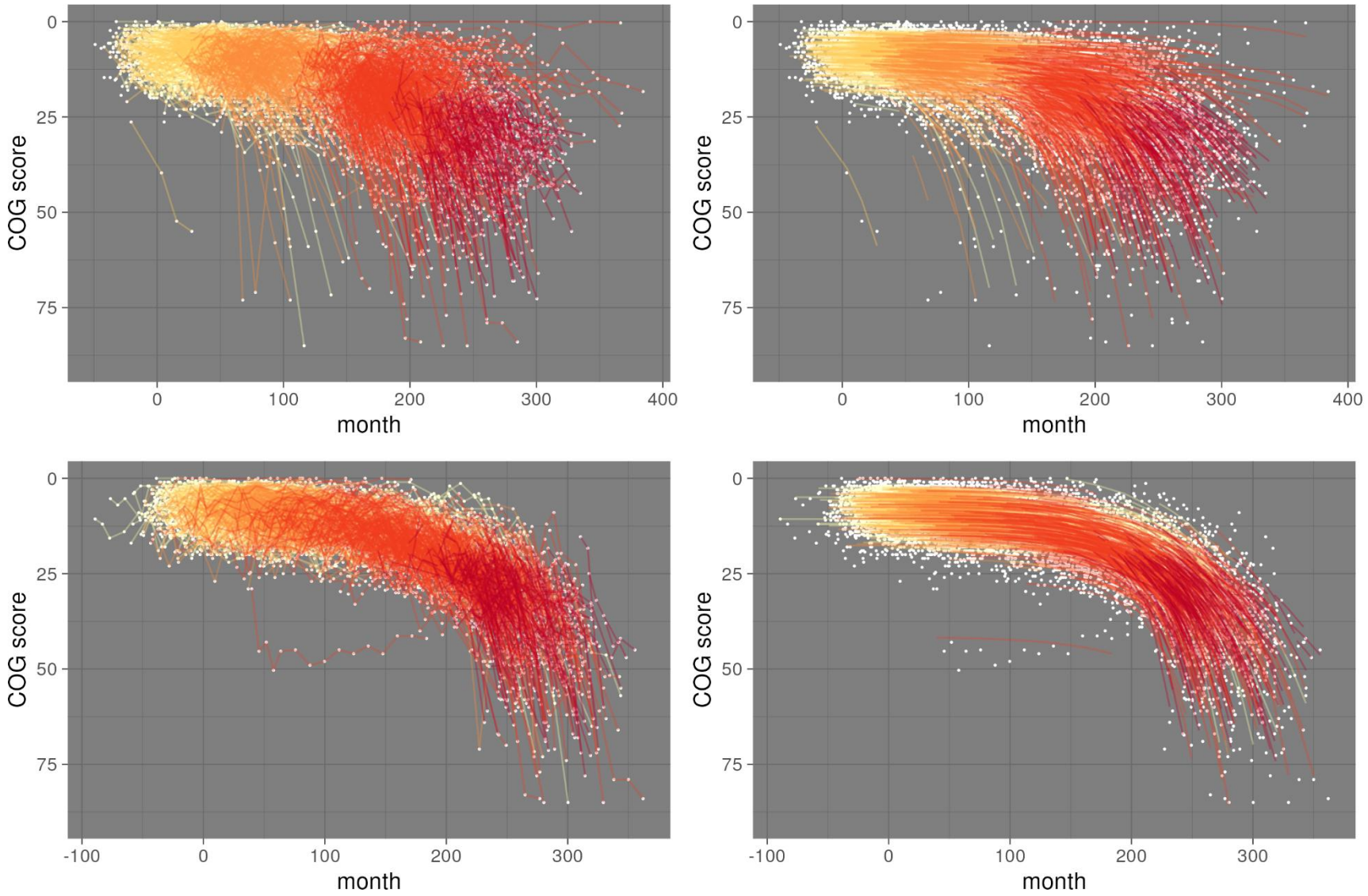
- Random intercepts model, but with a random x-intercept as well as a random y-intercept
- Each patient's data moved up/down and across onto disease-time trajectory
- Trajectory fitted simultaneously
- $\mu_{ij} = \alpha_i + \exp\left(\beta\left(\theta_i + t_{ij}\right)\right)$
  α_i = y-shift, θ_i = time-shift



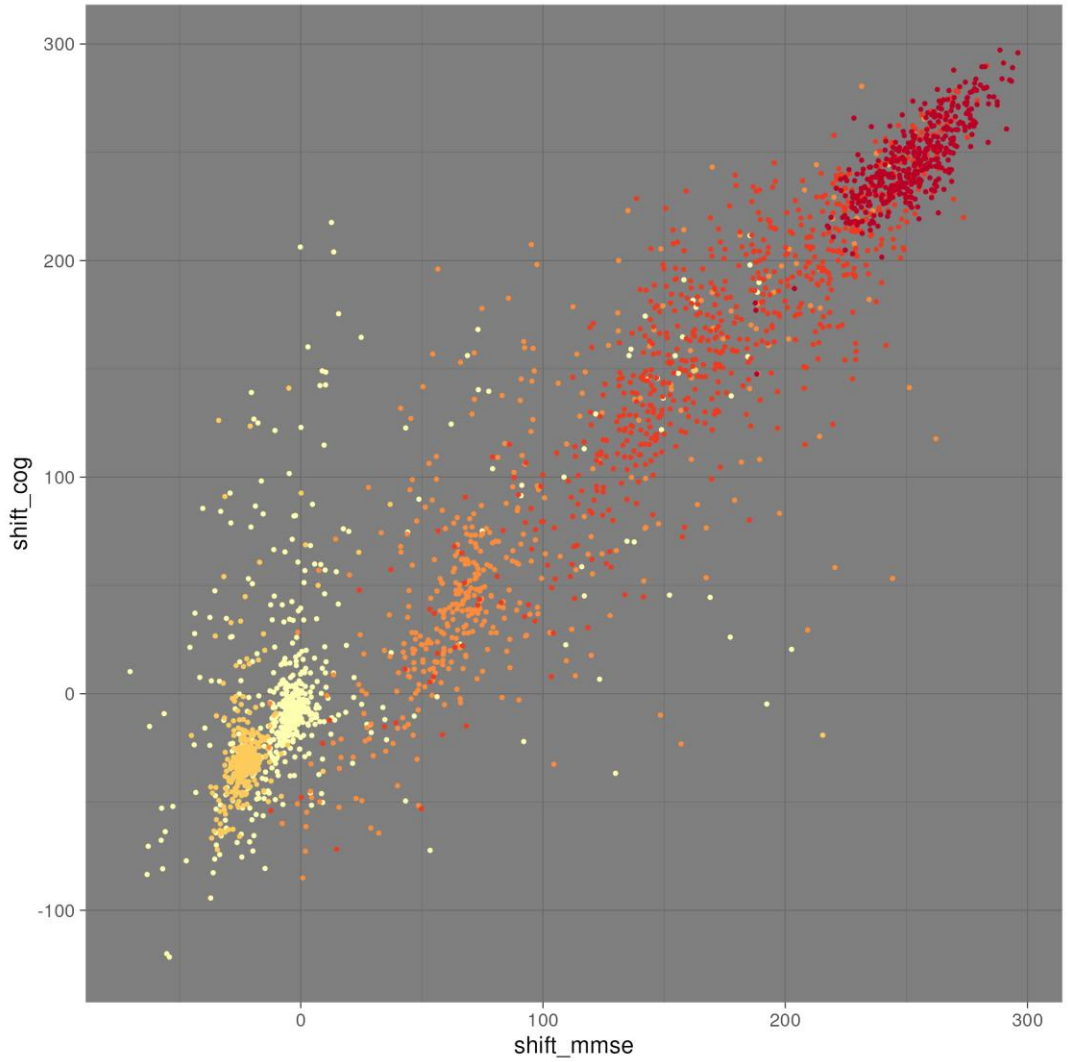
Outputs from Stan



Stan model: with baseline age, sex, education



Different/multiple endpoints: ADAS-Cog 13 vs MMSE

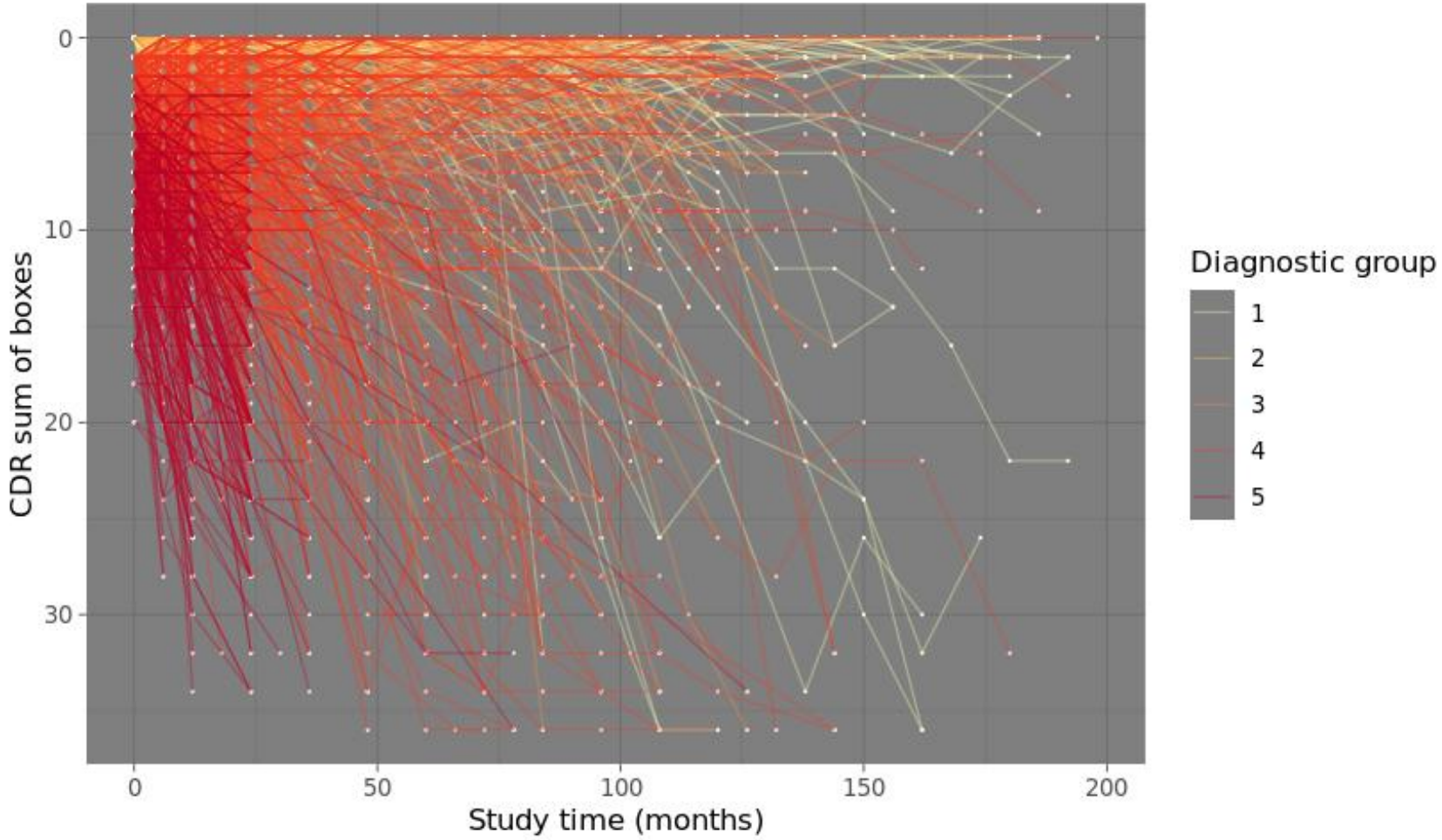


CDR Sum of Boxes (CDR-SB) = sum over 6 domains…

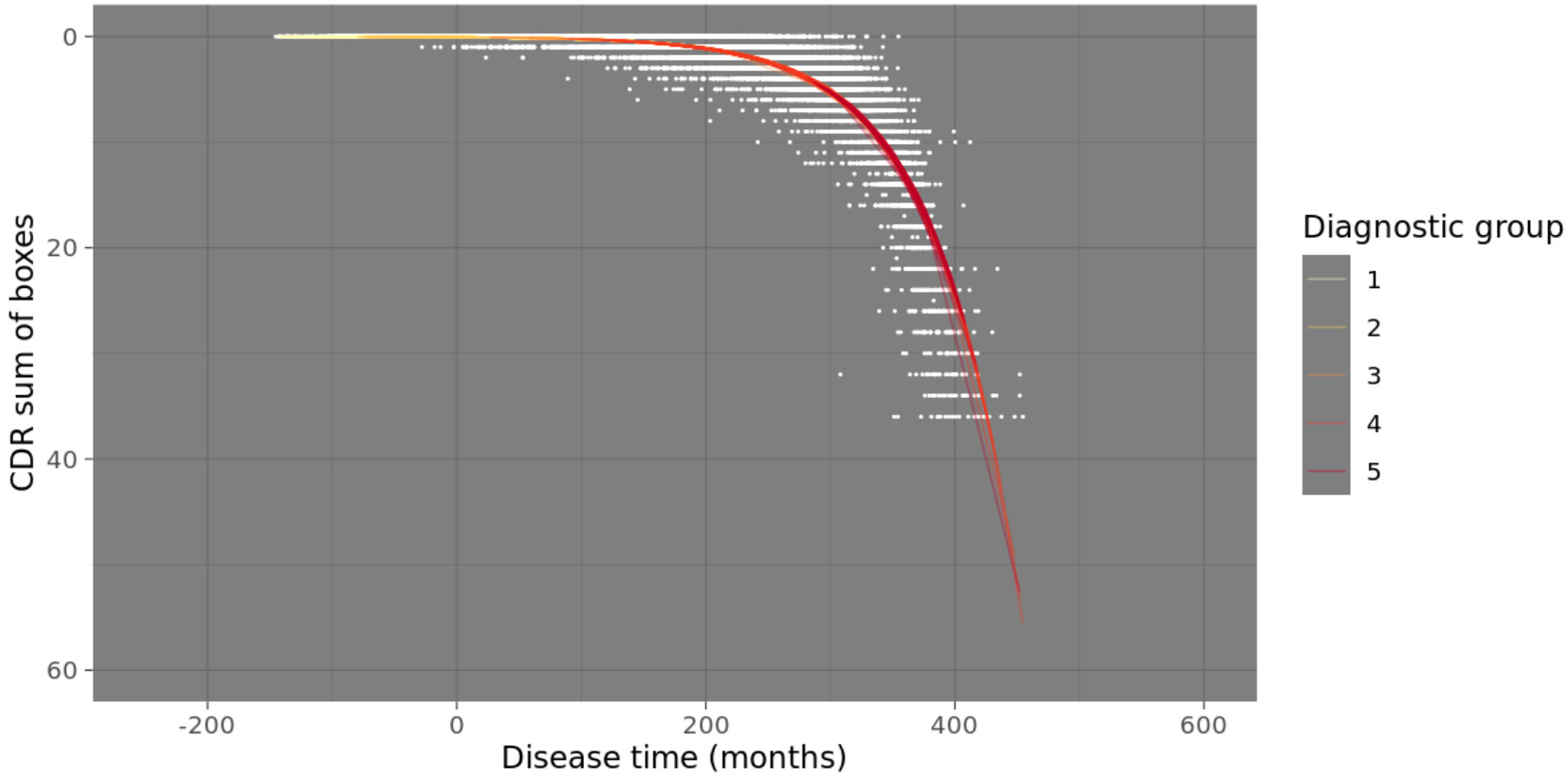
CDR Sum of Boxes	CDR Function	Community affairs
		Home and hobbies
		Personal care
	CDR Cognition	Memory
		Orientation
		Judgement/Problem-solving

- Total score between 0 and 18 with many zeros + half-integers
- Propose doubling scores and fitting negative-binomial model
- Start with Total Score; then fit Function + Cognition; and so on…

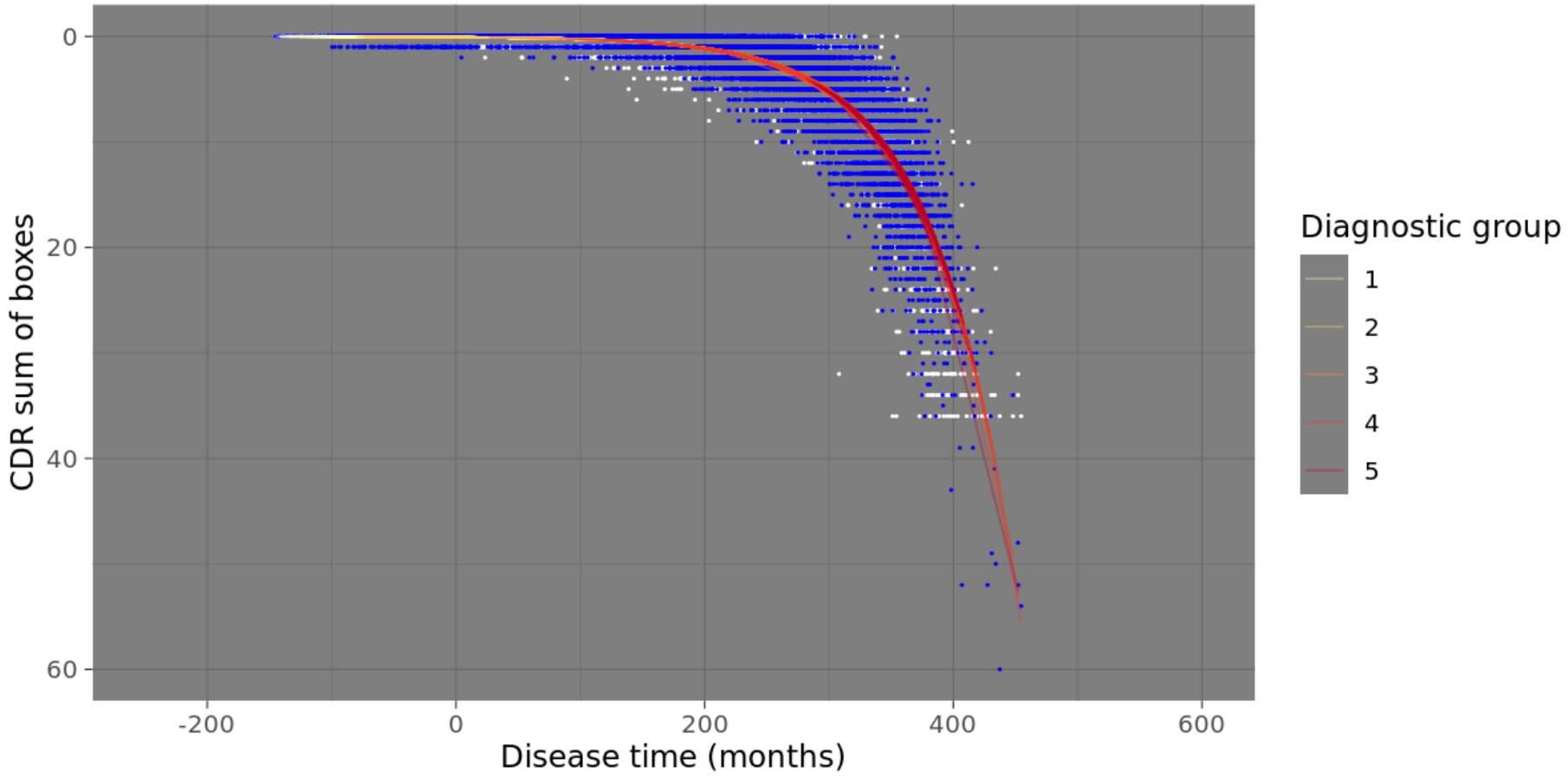
CDR-SB Total Score (x2)



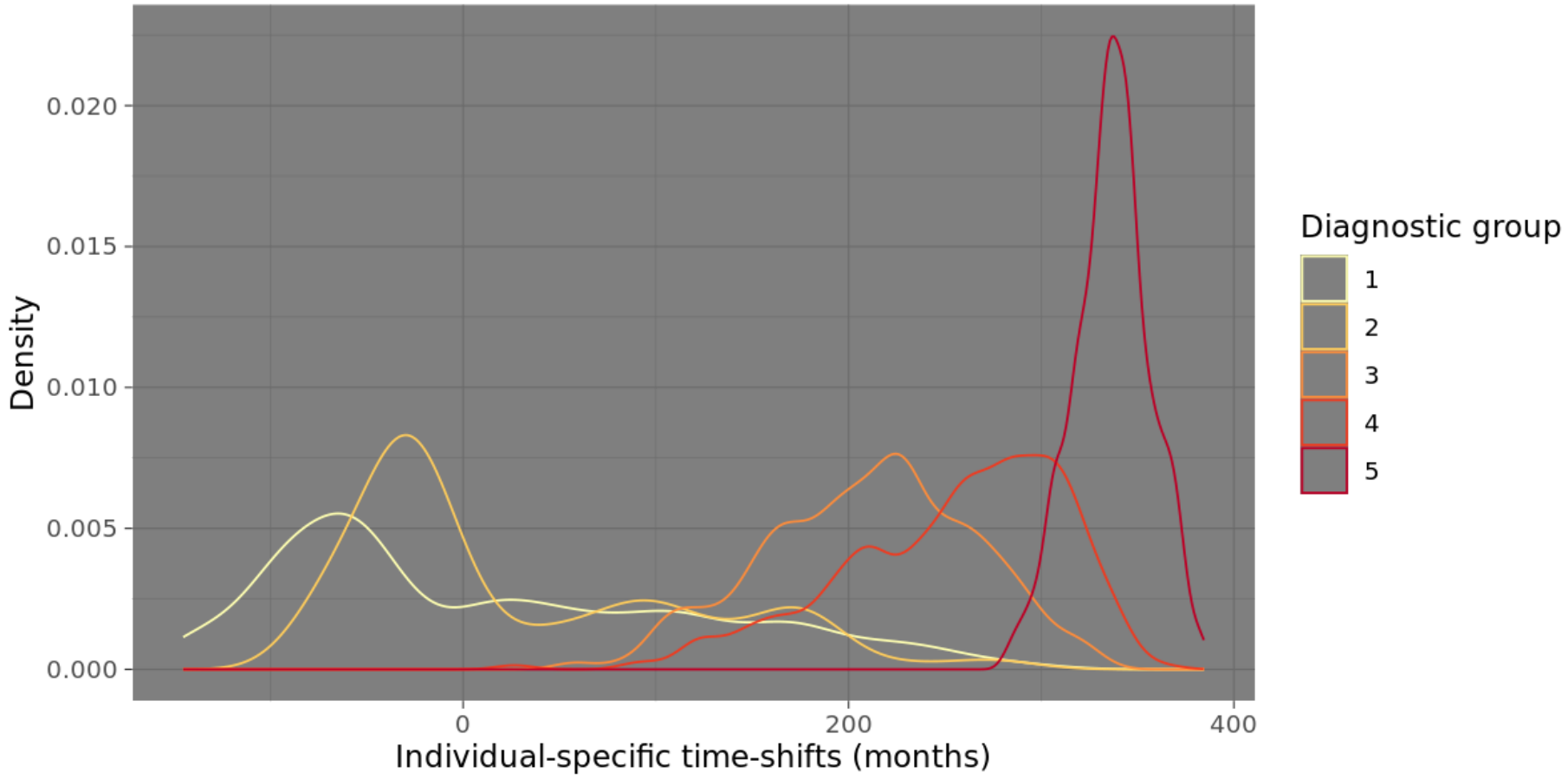
CDR-SB Total Score



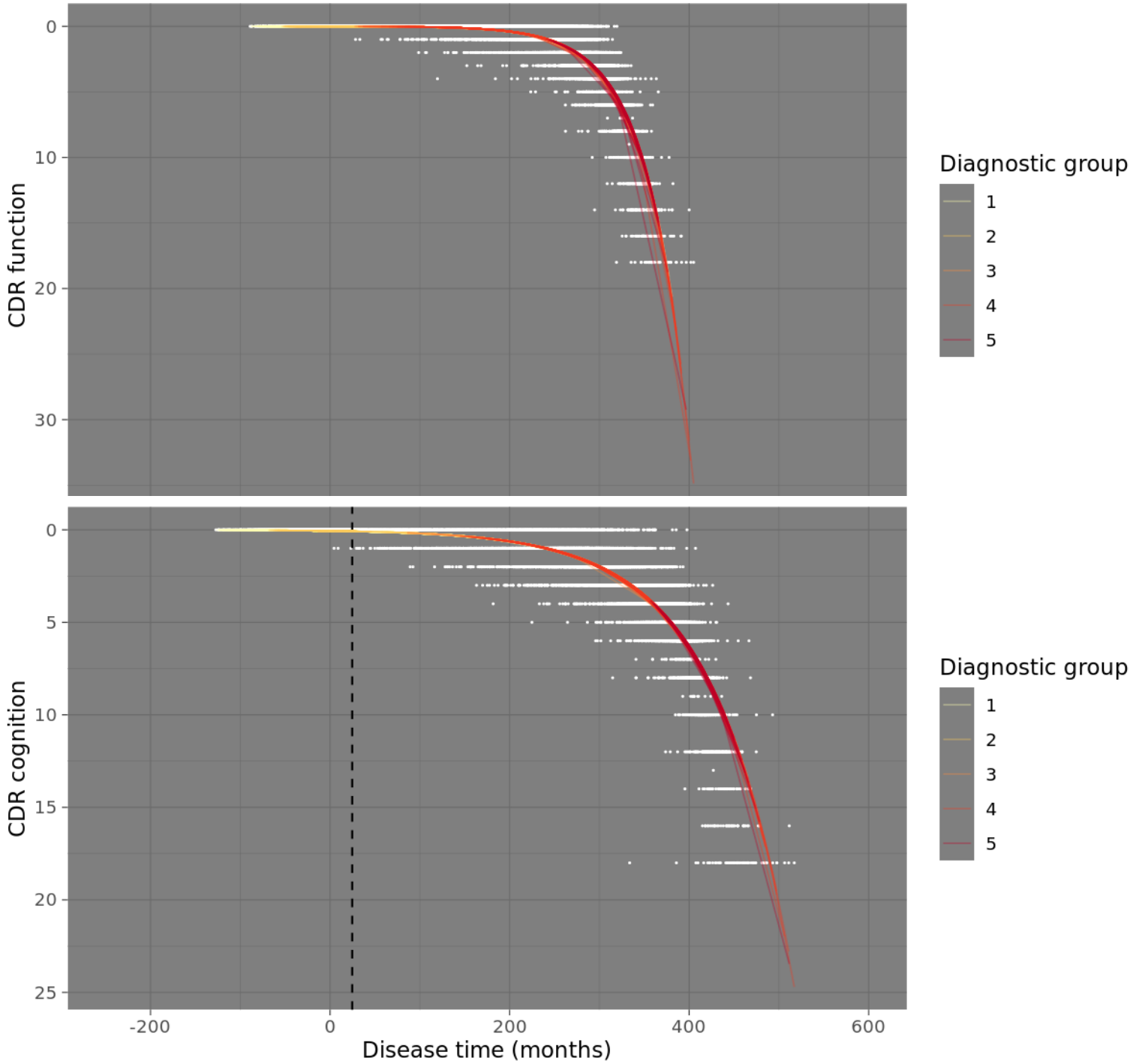
CDR-SB Total + simulated “noise”



CDR-SB Heterogeneity in individual-level time-shifts

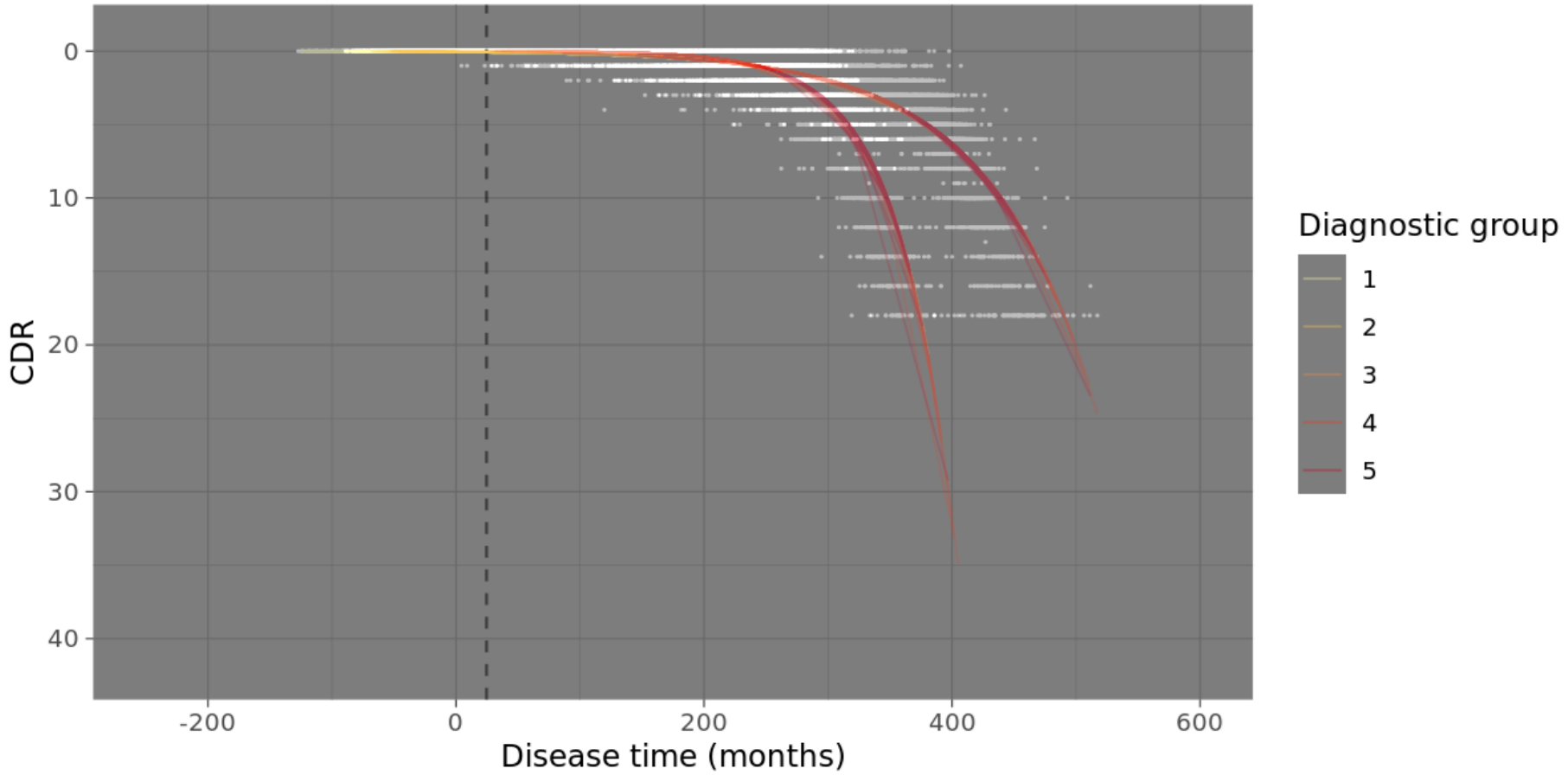


CDR-SB Function vs Cognition
Separate models



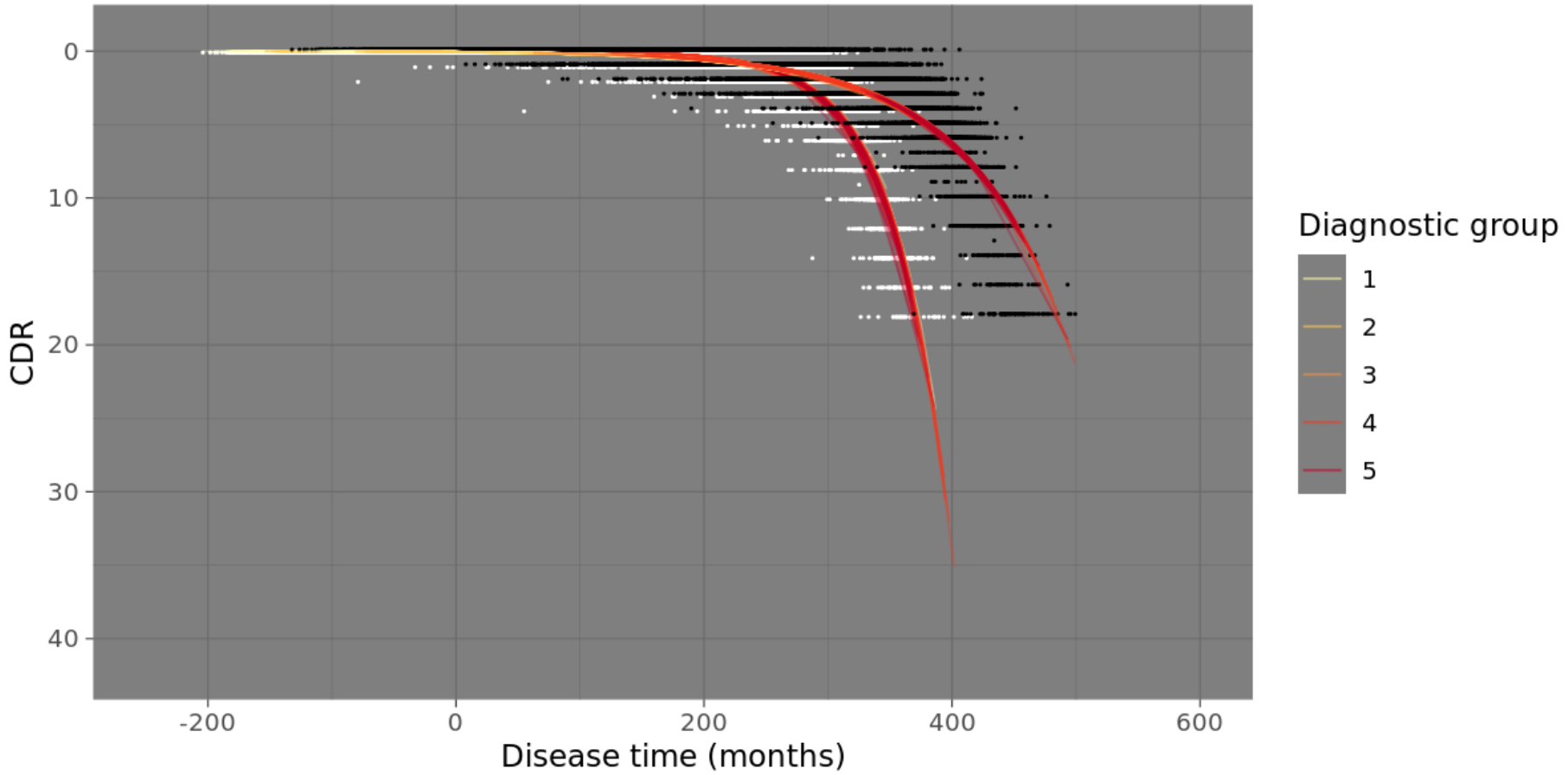
CDR-SB Function vs Cognition superimposed

Can this be modelled using a common “shift” random effect?



CDR-SB Function vs Cognition: Joint model

Cognition shift = linear transform of Function shift [$\theta_{Ci} = \eta + \lambda\theta_{Fi}$]



Overview

- Summary
  - Model synthesises all data into coherent whole: “population view” of full time-course
  - Understand/simulate rate of decline for different sub-populations, at different stages of disease
- Ongoing/future work [short term]
  - Build joint “natural history” model that reflects progression in a number of domains, and accounts for differences in the timing of progression within those domains
  - Identify patient characteristics (e.g. age, sex, p-tau217) that explain heterogeneity
    - Variable selection / ML model
    - Enables prediction of disease-age (DA) in trial populations
  - Natural history model can serve as “prior” for analysis of trial data
    - Joint nature allows evaluation of treatment effects beyond just cognition
  - Broader application: other data sources + disease areas (e.g. COPD)

Overview continued

- Ongoing/future work [longer term]
  - Simulate realistic trial data
  - Virtual populations/cohorts with characteristics (e.g. disease-stage, covariates) of interest
    - Refine inclusion criteria
  - Integrate/align with other models, e.g. enrollment, disease-state, QSP
    - End-to-end simulation
  - Hypothesise longitudinal treatment effects:
    - Explore how timing of treatment (and/or the observation schedule) and the type of trial population impact on the identification of treatment effects
    - Different ways of measuring treatment effects, e.g. residence time

Thank you for your attention

- $y_{dij} \sim D(\mu_{dij}, s_d)$, i = patient, j = visit, d = domain;
- $\mu_{dij} = \alpha_{di} + \exp\left(\beta_d(\delta_d + \theta_{di} + t_{dij})\right)$, α_{di} = y-shift, θ_{di} = time-shift;
- $\theta_{Fi} \sim N(\gamma_g, \phi_g^2)$, g = diagnosis group;
- $\theta_{Ci} = \eta + \lambda\theta_{Fi}$ [for example]