

Multi-Endpoint Item-Response Modelling for Integrating Several Endpoints in Alzheimer's Disease Drug Trials

Camille Vong (1), Shuying Yang (2), Massimiliano Germani (3), Chao Chen (2)

Clinical Pharmacology Modelling and Simulation, (1) GSK, Baar, Switzerland, (2) GSK, London, UK, and (3) GSK, Tres Cantos, Spain.



As a proof-of-concept, the approach of integrating several clinical scales to inform the underlying latent constructs using MIRMs increased power for detecting treatment effect. The resulting latent variable can serve as surrogates for clinical endpoints with high signal-to-noise ratio.

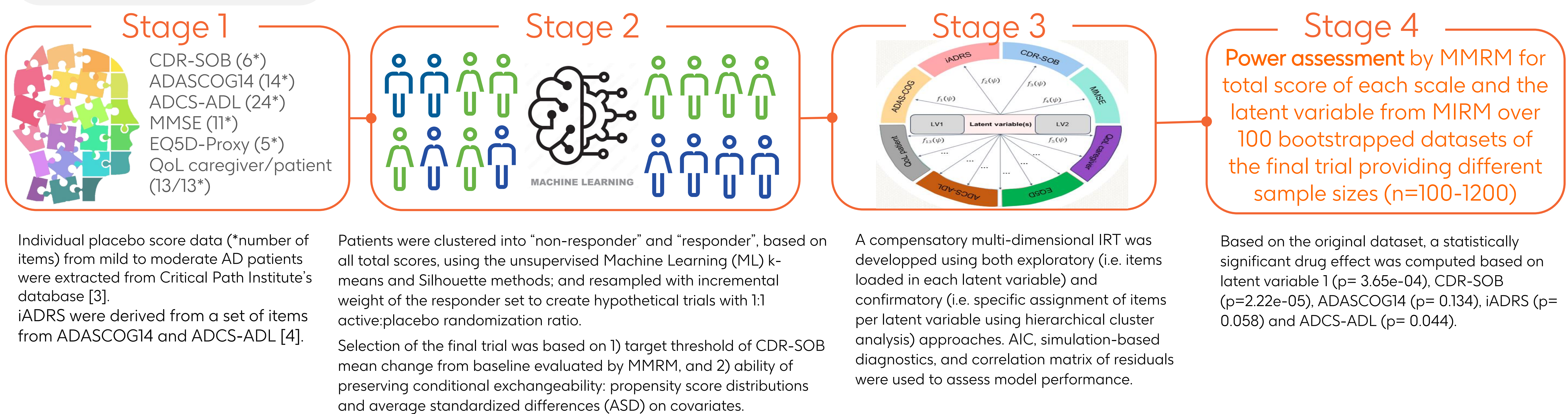
Background

- Several rating scales are used in Alzheimer's Disease (AD) to describe the heterogeneity of the disorder as a collection of cognition, daily function, behaviour and quality of life assessments.
- They remain nonetheless an arduous and lengthy procedure for both physician and patient/caregiver [1]. These scales have demonstrated low efficacy signal to noise ratio [2], and high redundancy.

Aims

The objectives of this analysis were to build a multi-endpoint item-response model (MIRM), integrating data from different clinical scales. The performance of the MIRM was tested within a conceptual randomized trial framework, created entirely from pooled placebo data.

Methods



Results

A confirmatory 2-latent MIRM (86 items and 480 parameters) with a correlation of 0.221 performed better ($\Delta AIC = -8412.7$) than the 1-latent MIRM. LV2 did not show difference between placebo and active, hence was dropped from the power assessment.

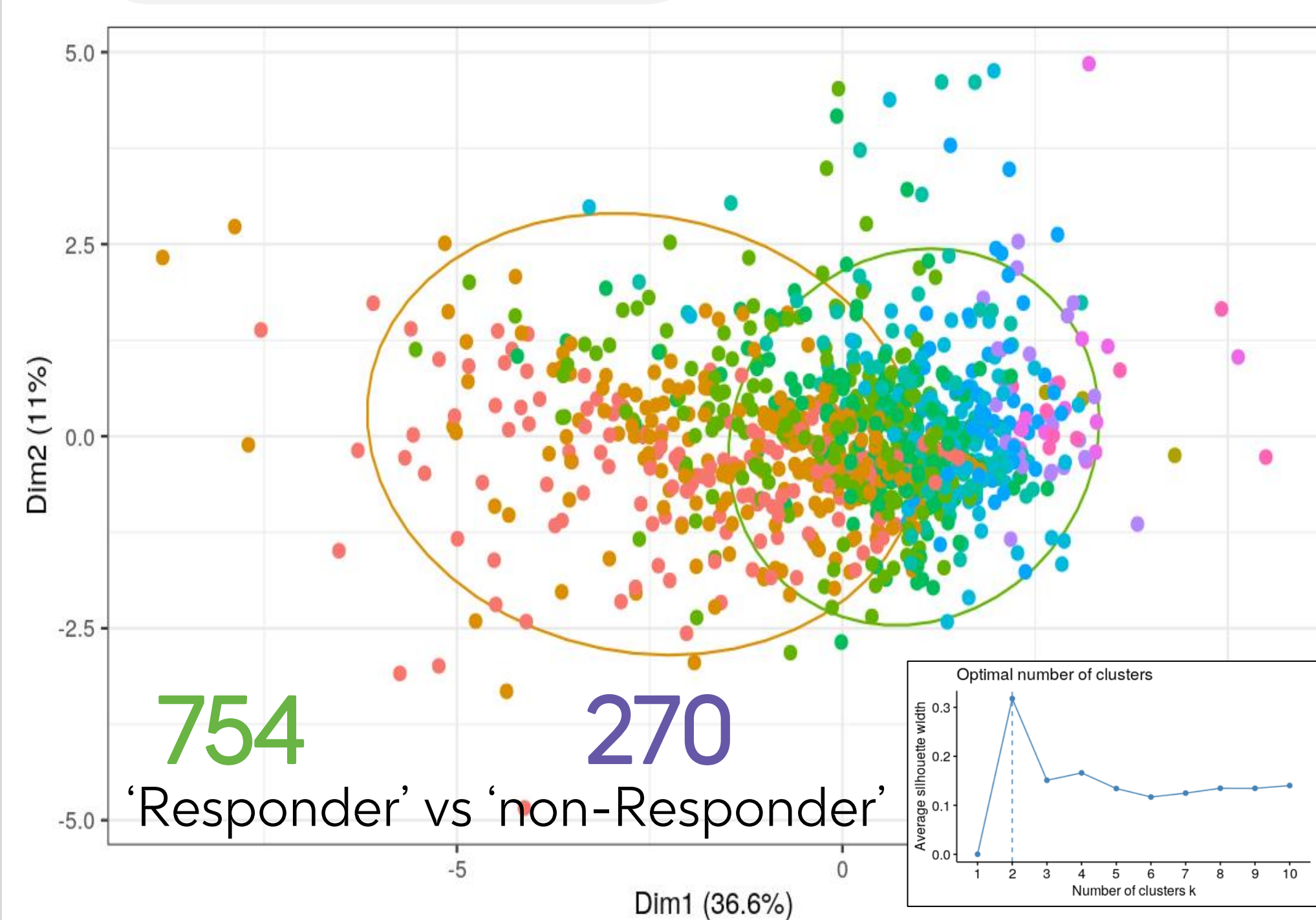


Figure 1: Classification of 'non-responder' and 'responder' patients based on k-means, guided by average silhouette.

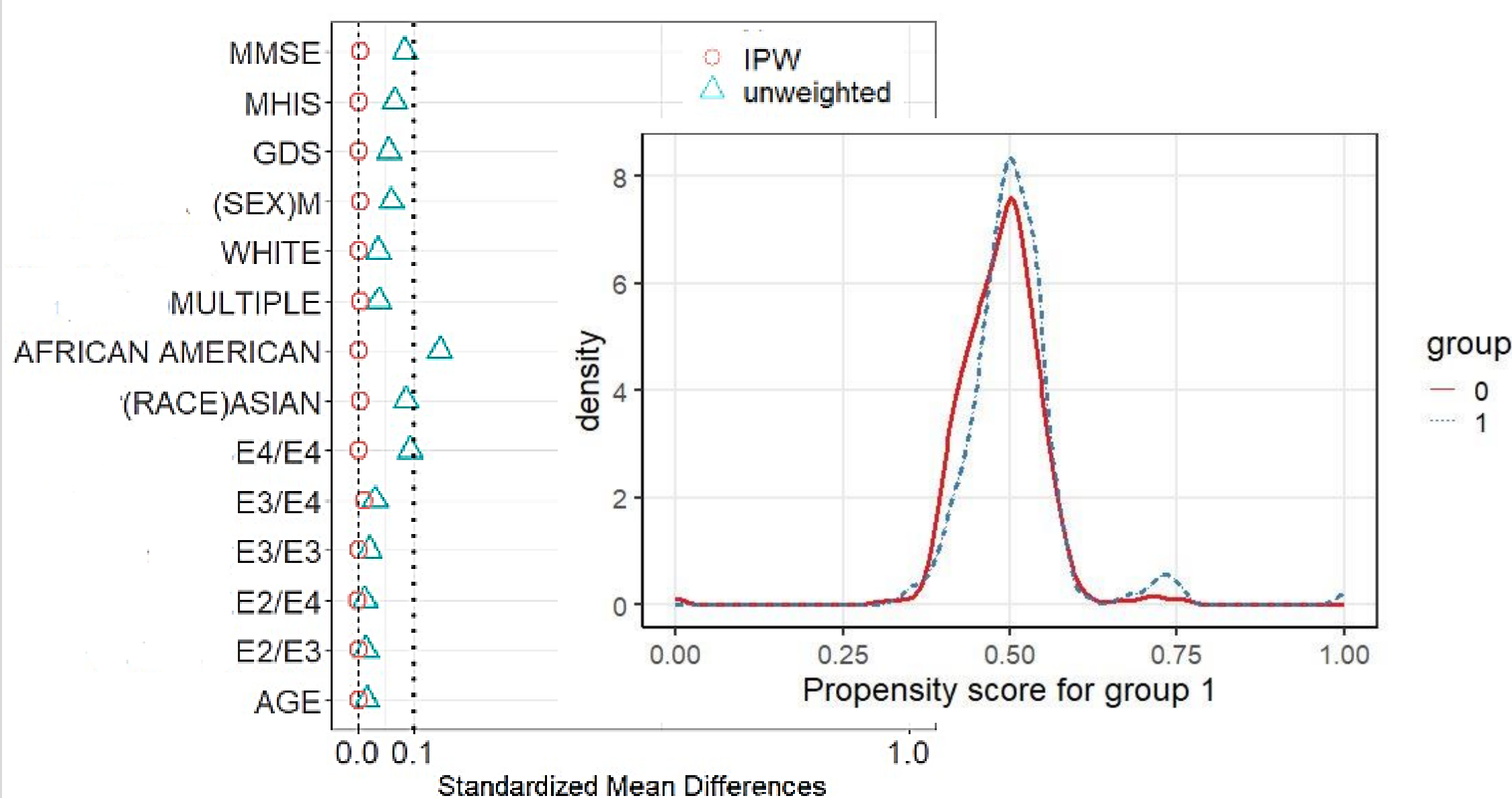


Figure 2: Satisfactory conditional exchangeability evaluated by overlapped propensity score distributions and ASD scores <0.1 on covariates.

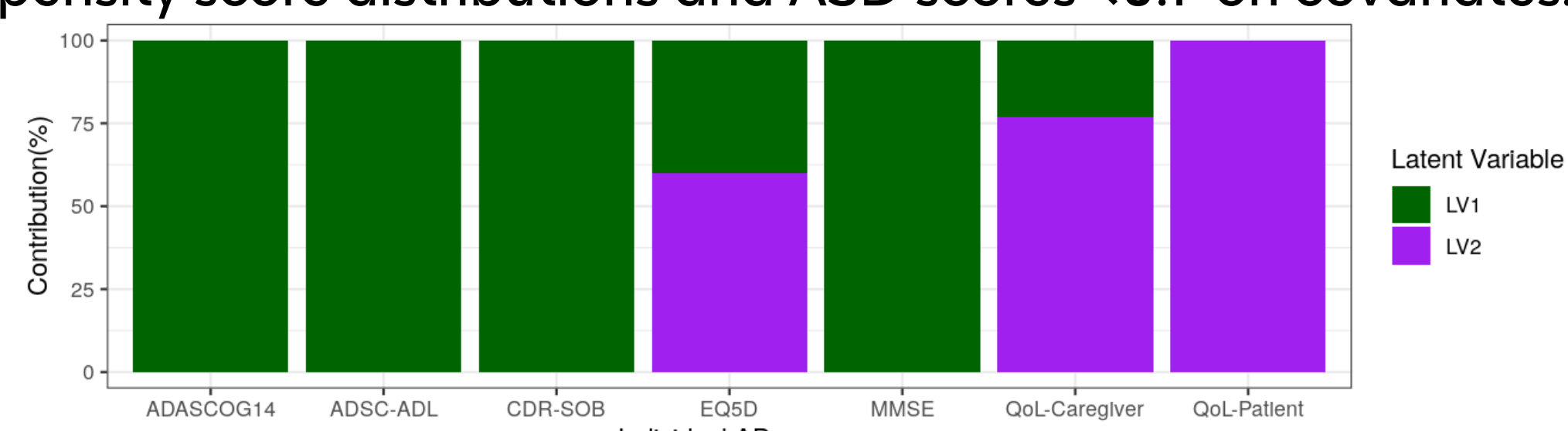


Figure 3: Contribution of scales to the 2 latent variables revealed by factor loadings.

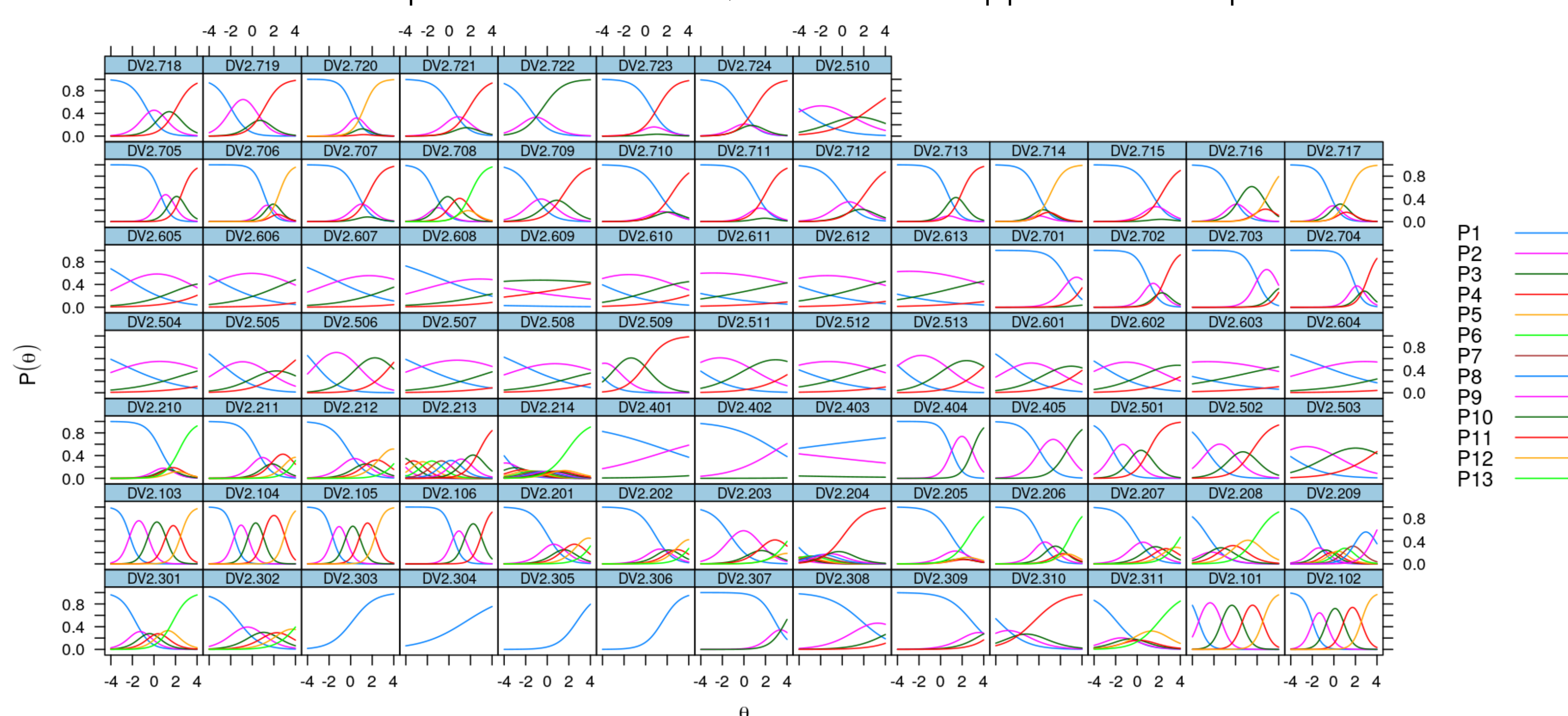


Figure 4: One-dimension view of item characteristic curves showing the individual probability of obtaining scores in each category for CDR-SOB items, labelled 100s, ADASCOG14 - 200s, MMSE - 300s, EQ5D - 400s, QoL - 500/600s (patient/caregiver) and ADCS-ADL - 700s.

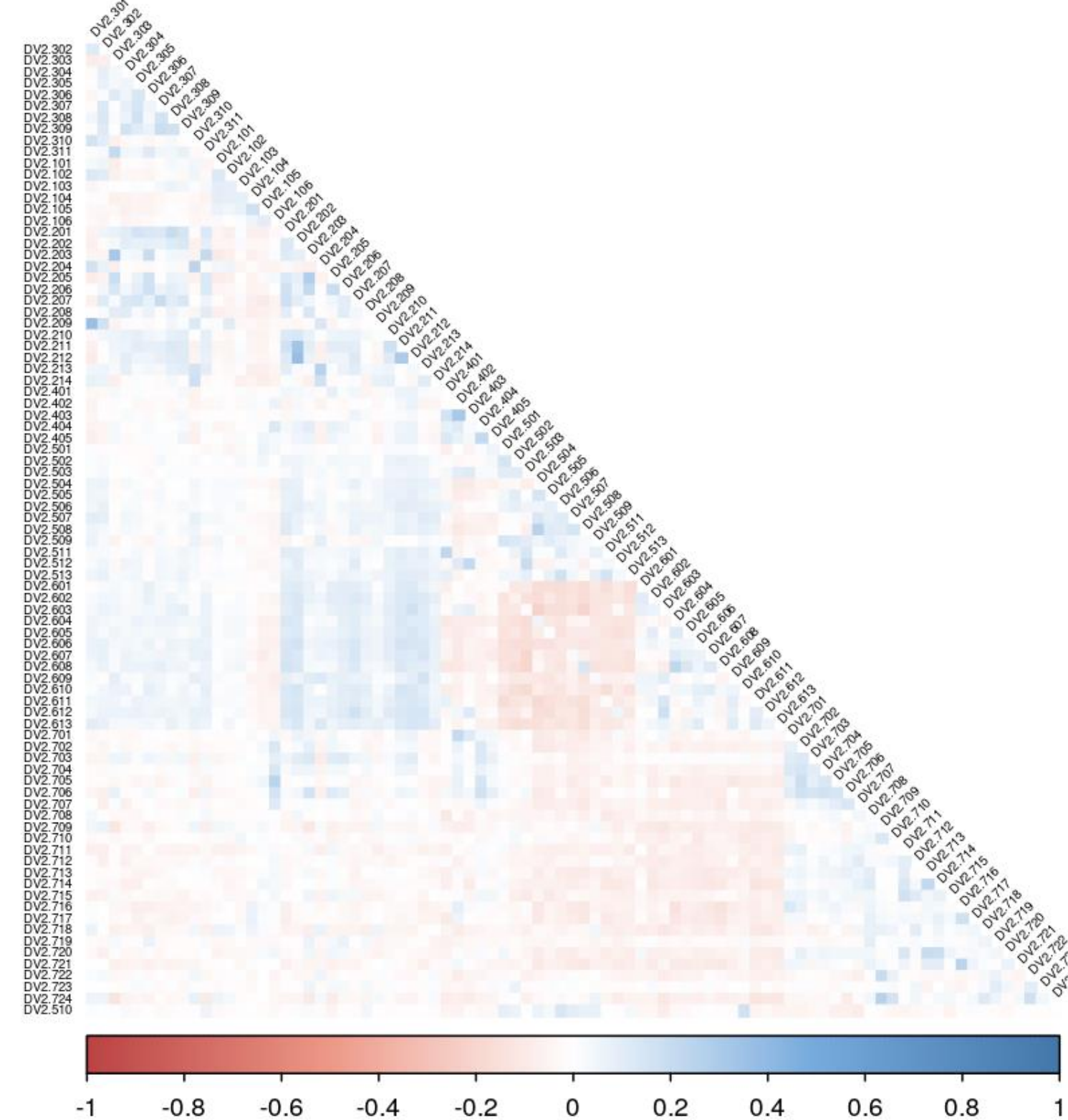


Figure 5: Residual correlation plot between all pairs of items of all AD scores. General observed correlations are mostly between -0.5 and 0.5.

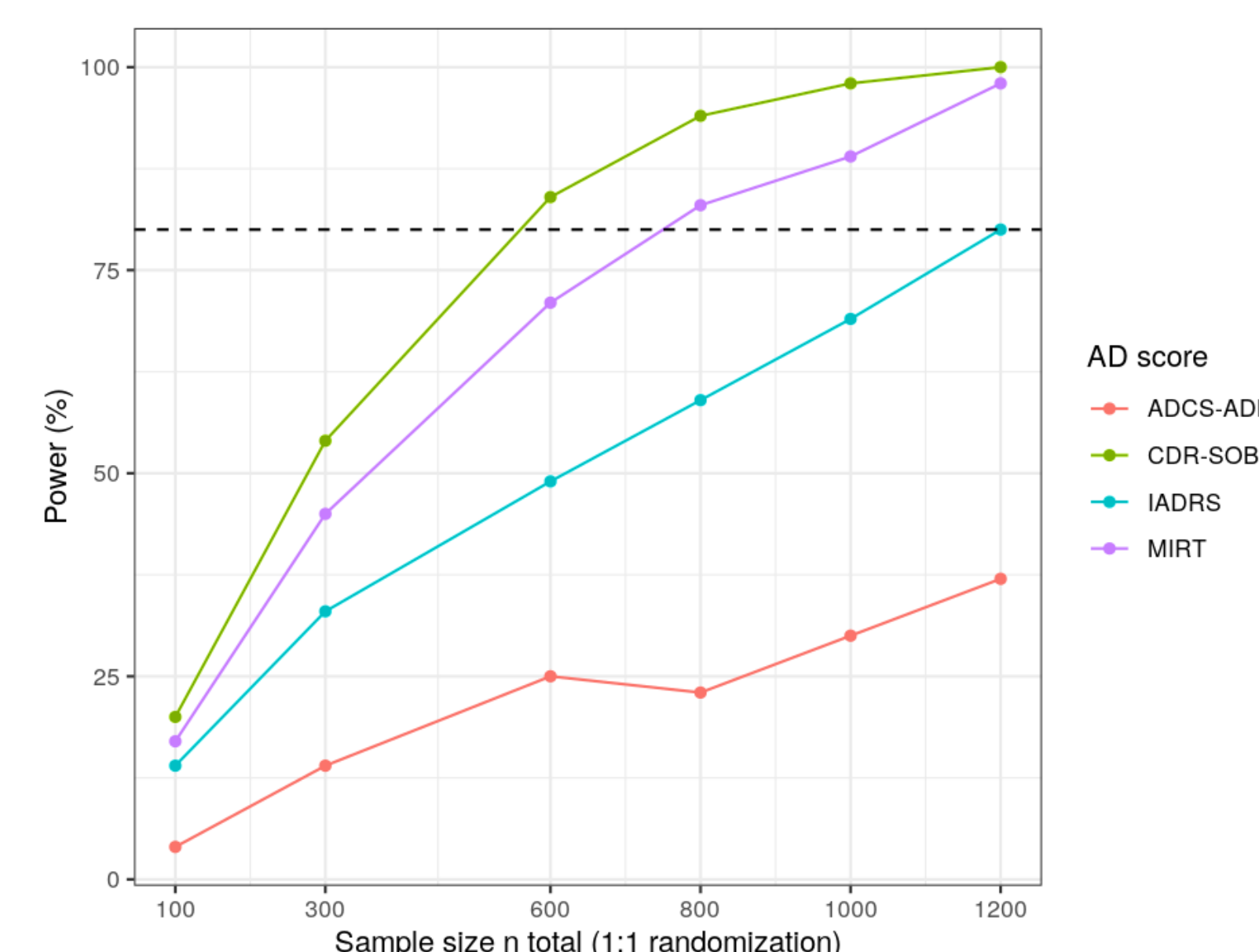


Figure 6: The final MIRM showed that with 800 subjects the latent variable yielded 83% power to detect a treatment effect, while CDR-SOB, iADRS or ADCS-ADL alone resulted in 94%, 59% and 25% power, respectively.