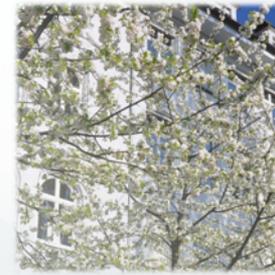


Max Planck Institute
of Psychiatry

German Research Institute of Psychiatry

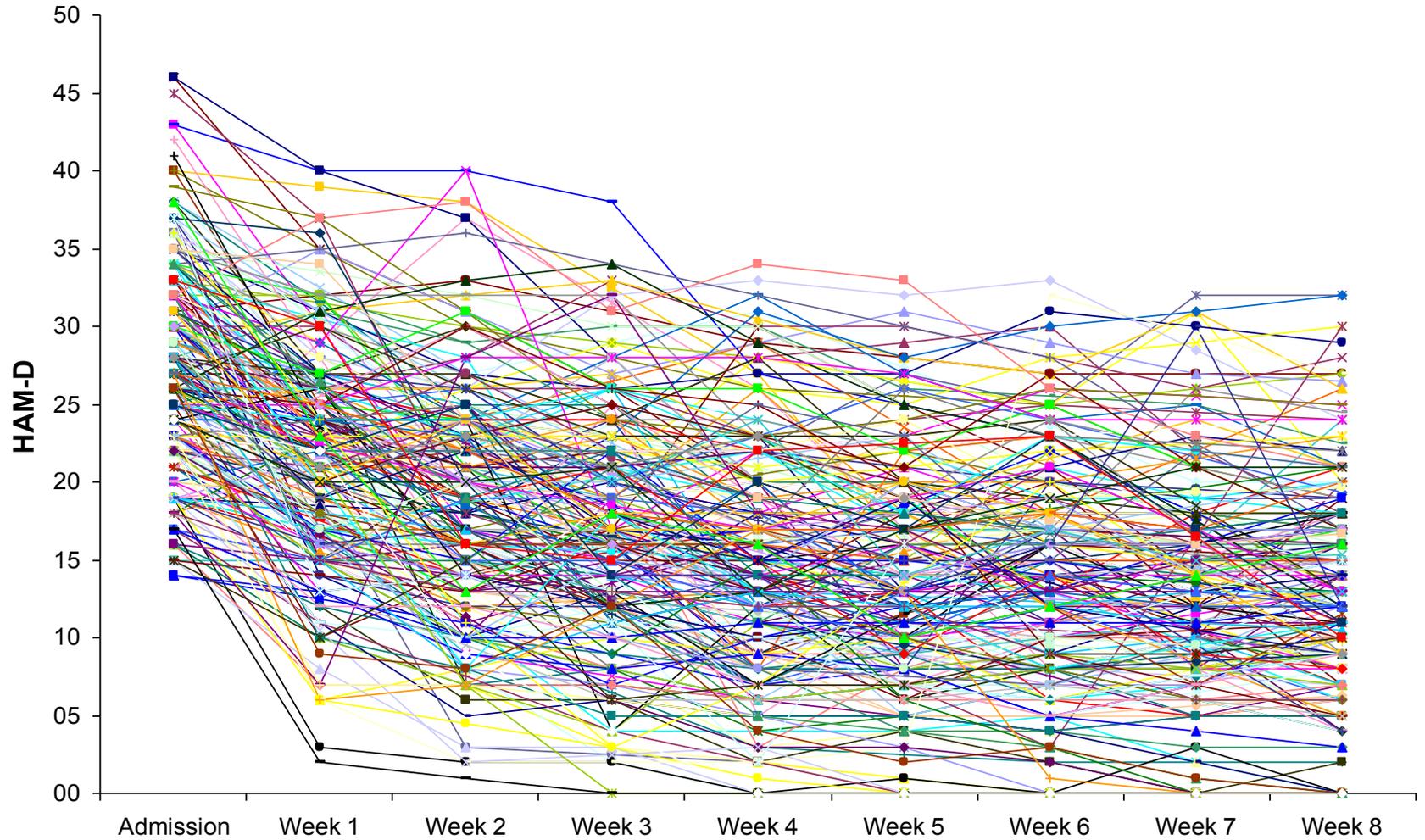


Deciphering subtypes in response to treatment in major depression - a longitudinal approach

Bertram Müller-Myhsok

bmm@psych.mpg.de

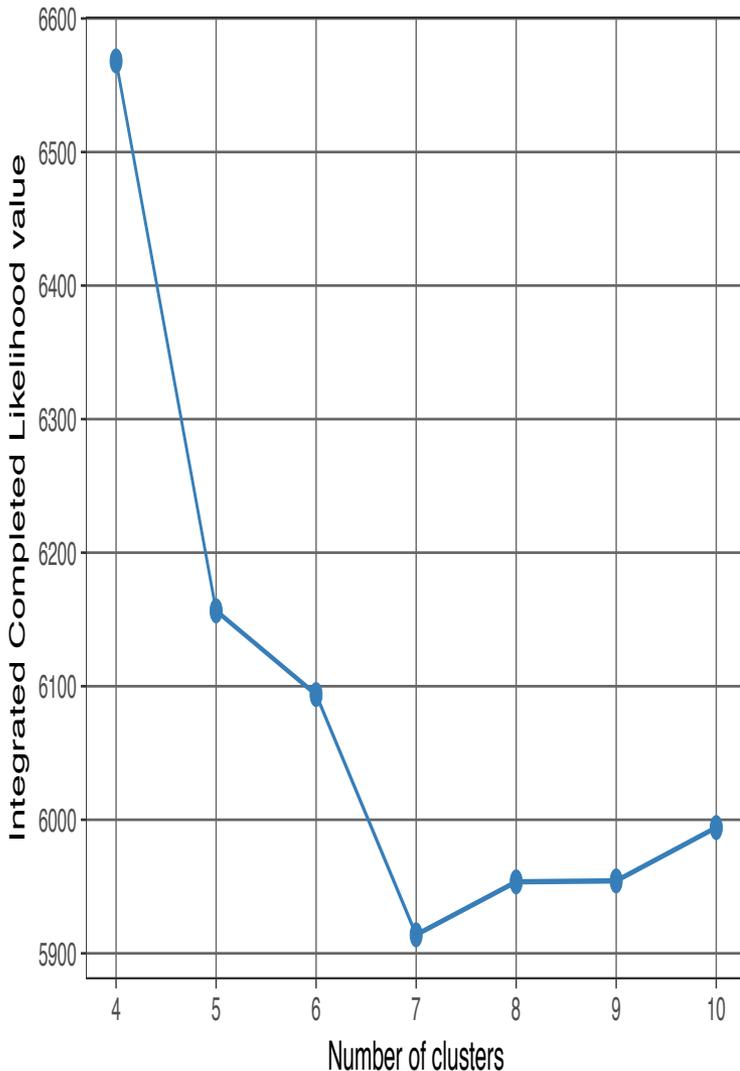
Heterogeneous Antidepressant Treatment Response Pattern



Short Papers

Assessing a Mixture Model for Clustering with the Integrated Completed Likelihood

Christophe Biernacki,
Gilles Celeux, and Gérard Govaert

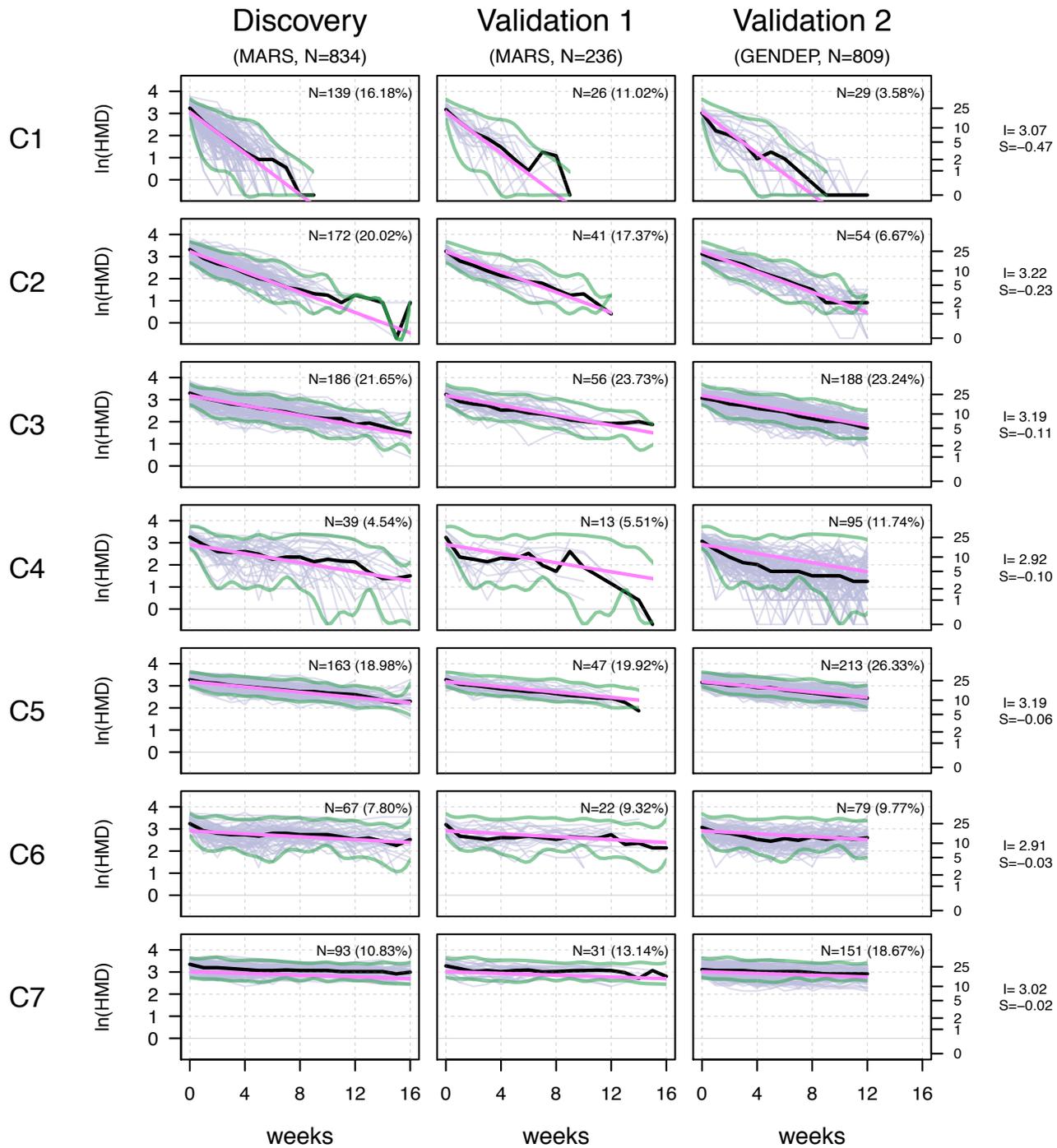


Abstract—We propose assessing a mixture model in a cluster analysis setting with the integrated completed likelihood. With this purpose, the observed data are assigned to unknown clusters using a maximum a posteriori operator. Then, the Integrated Completed Likelihood (ICL) is approximated using an *à la* Bayesian information criterion (BIC). Numerical experiments on simulated and real data of the resulting ICL criterion show that it performs well both for choosing a mixture model and a relevant number of clusters. In particular, ICL appears to be more robust than BIC to violation of some of the mixture model assumptions and it can select a number of clusters leading to a sensible partitioning of the data.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,
VOL. 22, NO. 7, JULY 2000

$$\text{ICL}(m, K) =$$

$$\log \mathbf{f}(\mathbf{x}, \hat{\mathbf{z}} \mid m, K, \hat{\theta}) - \frac{\nu_{m,K}}{2} \log n.$$



Supplemental Table S1. Baseline HAM-D and average HAM-D values per cluster (discovery sample)

Cluster label	Baseline HAM-D [mean (SD)]	Average HAM-D across time series ^a [mean (SD)]
C1 ^b	24.63 (6.08)	11.15 (3.74)
C2	26.35 (6.26)	14.15 (3.74)
C3	26.10 (6.14)	15.47 (3.95)
C4	26.22 (6.65)	12.46 (3.53)
C5	26.71 (4.86)	19.17 (4.07)
C6	25.10 (6.44)	16.23 (4.30)
C7	27.04 (6.03)	21.80 (4.56)
ANOVA (df=6)	$p=9.446 \times 10^{-3}$	$p=4.022 \times 10^{-116}$
Linear correlation ^c	$r=0.09, p=2.497 \times 10^{-3}$	$r=0.57, p=8.271 \times 10^{-76}$

^a Average across all available HAM-D values of the time series until discharge.

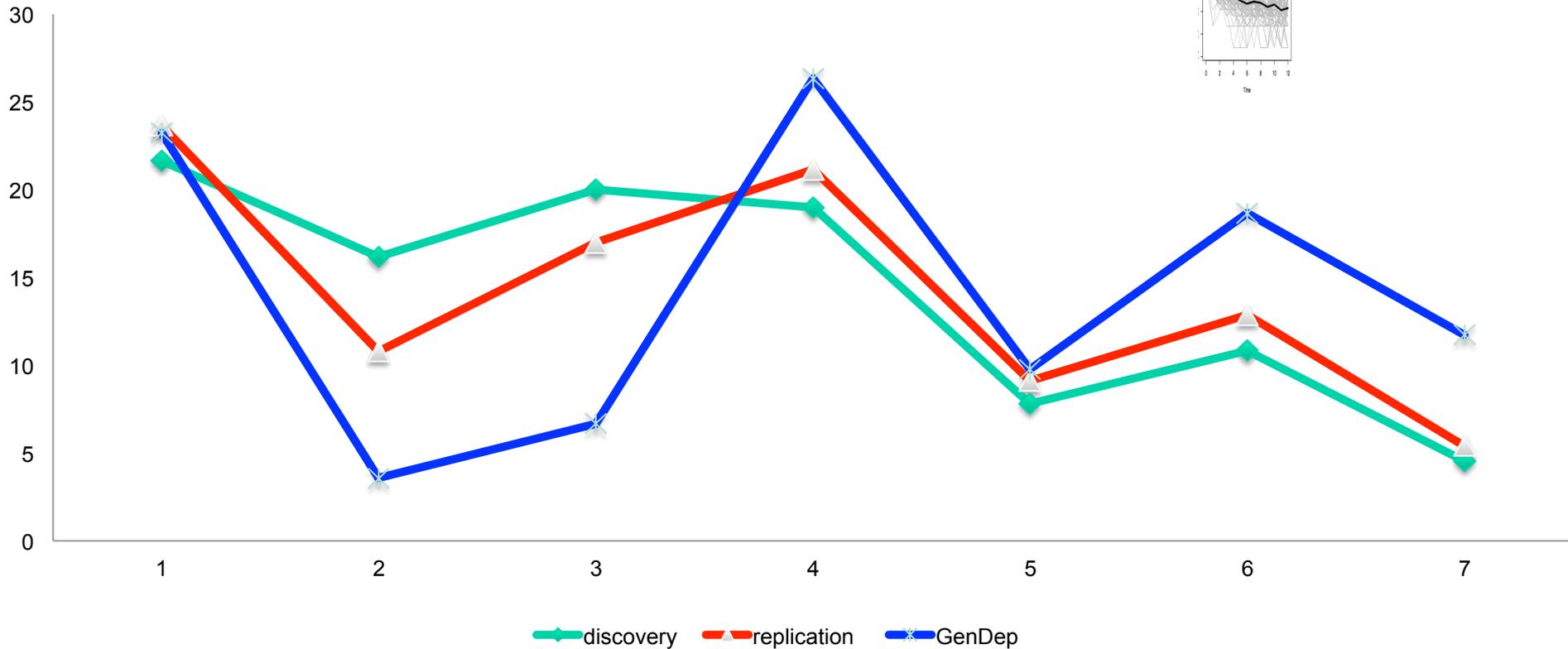
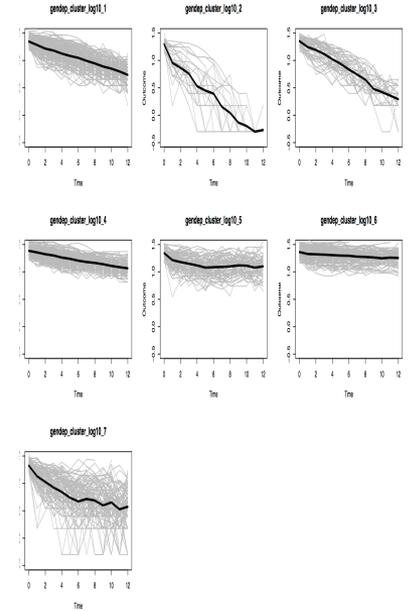
^b Sorting is by increasing cluster-derived slope, as in **Figure 1**.

^c Correlation between cluster-derived slope and individual baseline HAM-D values (middle column) and individual average HAM-D across time (right column).

Distribution of clusters

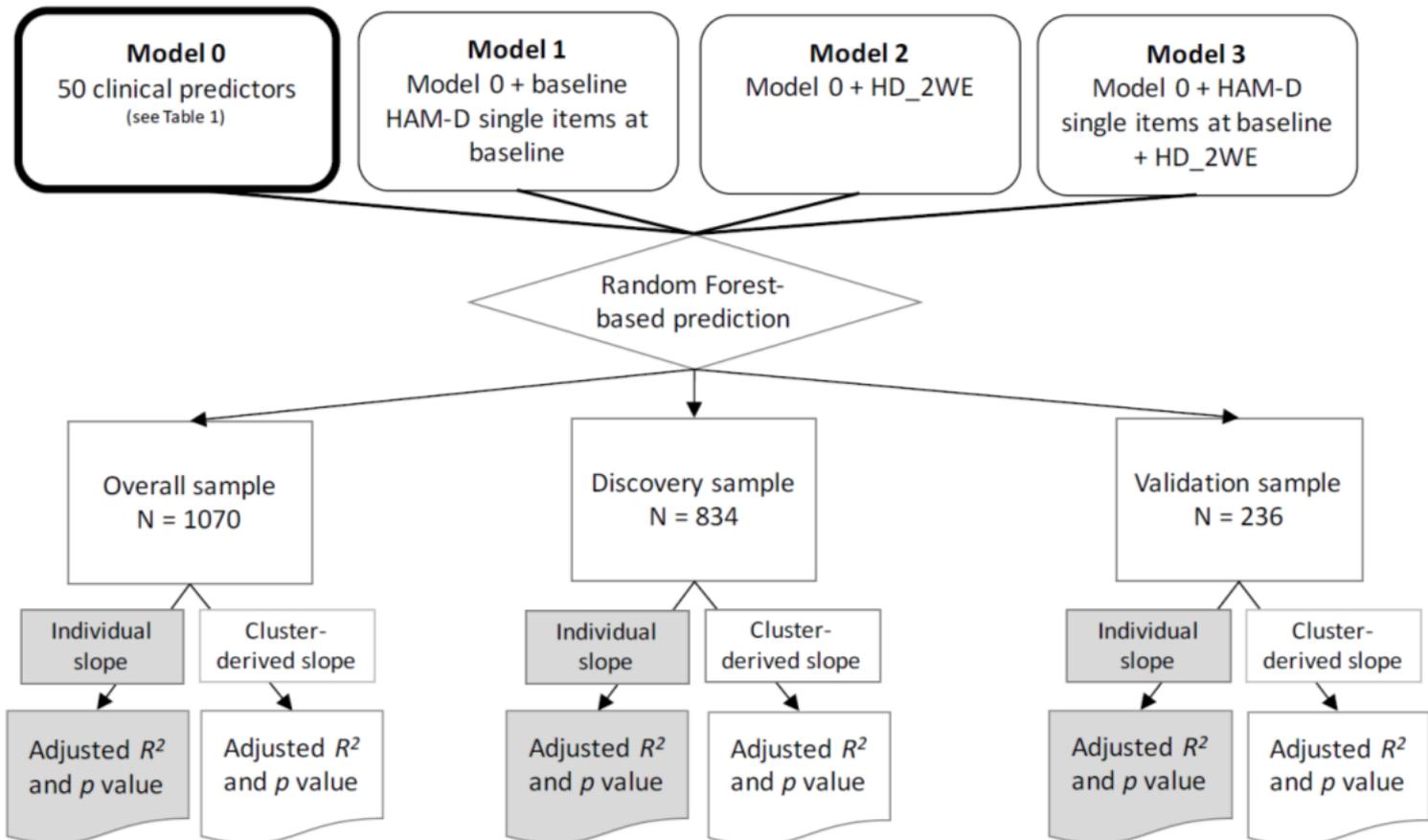
	C1	C2	C3	C4	C5	C6	C7
discovery	186	139	172	163	67	93	39
replication	57	26	41	51	22	31	13
GenDep	188	29	54	213	79	151	95

GenDep

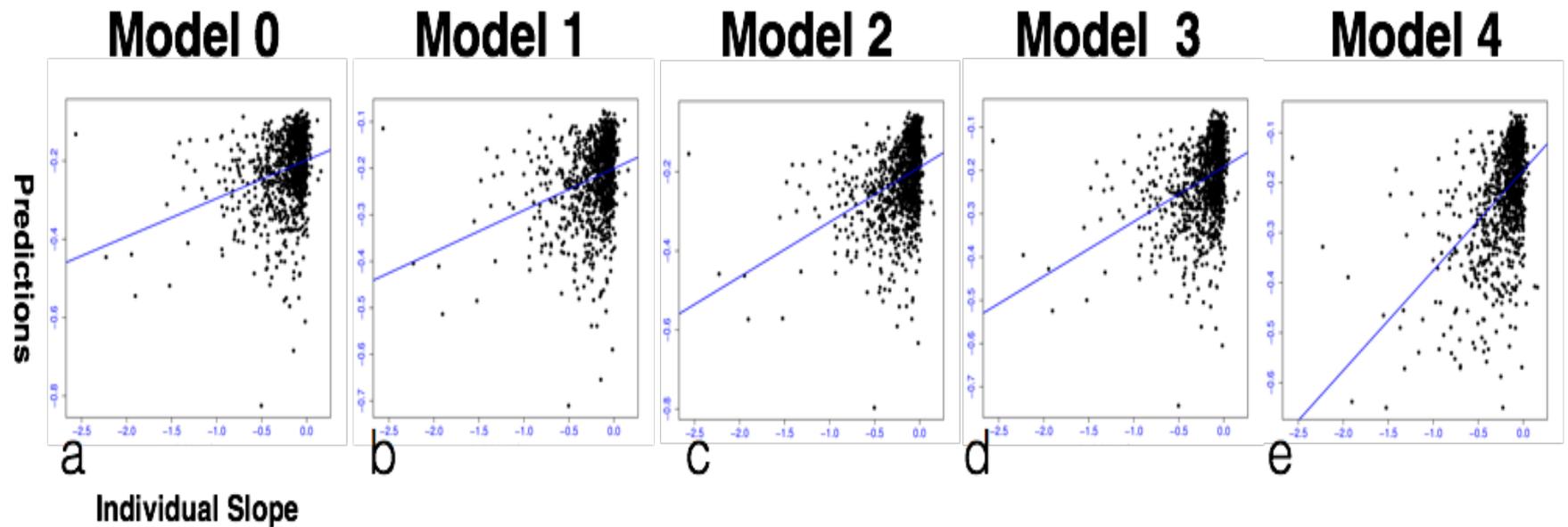


Prediction - verify clusters by relating back to clinical variables

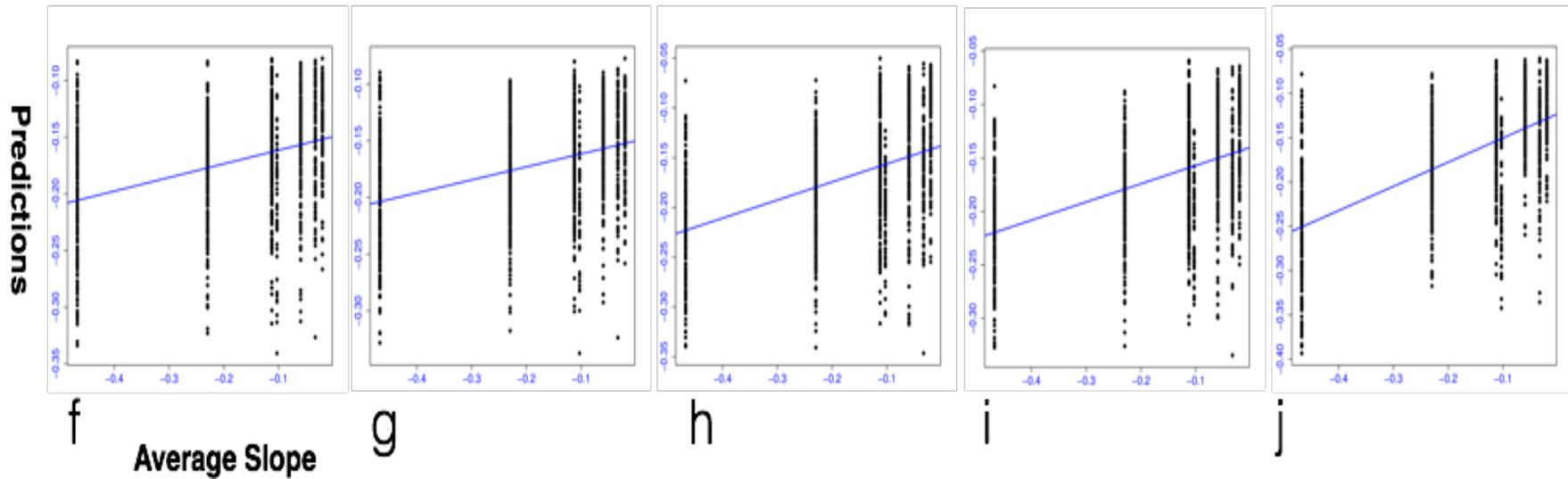
- **Prediction algorithm**
 - **Random forest**
 - **Clinical variables (n=80)**
 - **EPQ**
 - **TPQ**
 - **Life events**
 - **Symptom checklist**
 - **HAMD at week 0**
 - **[HAMD at week 2]**
 - **...**



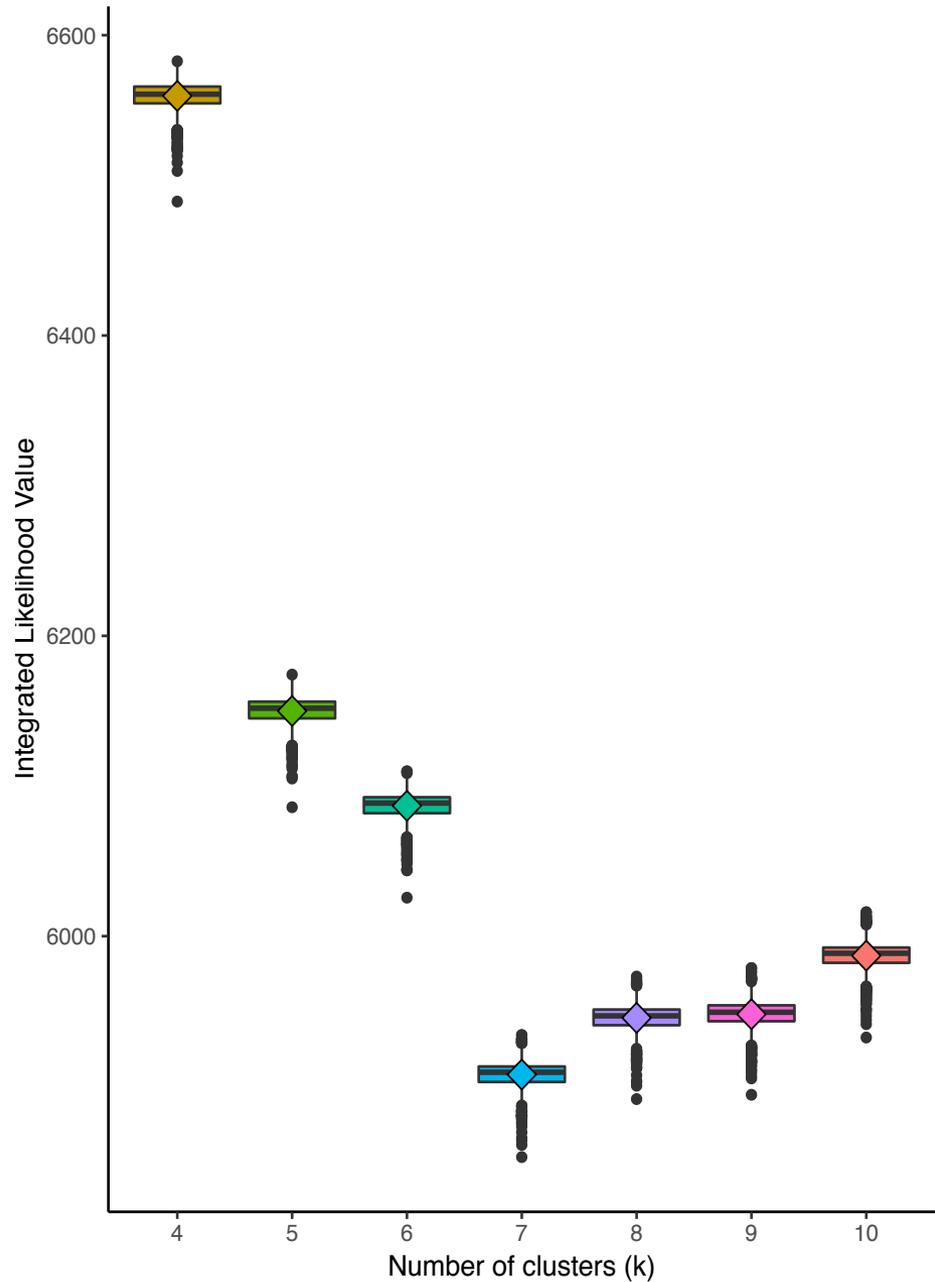
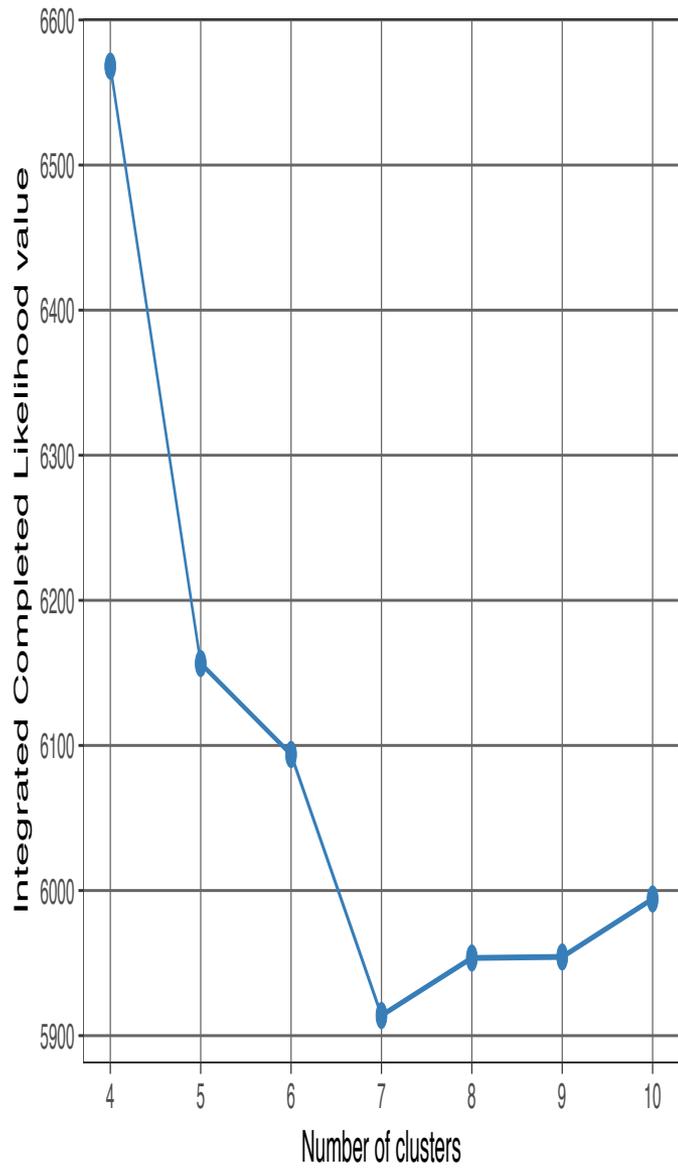
Using Individual Slope

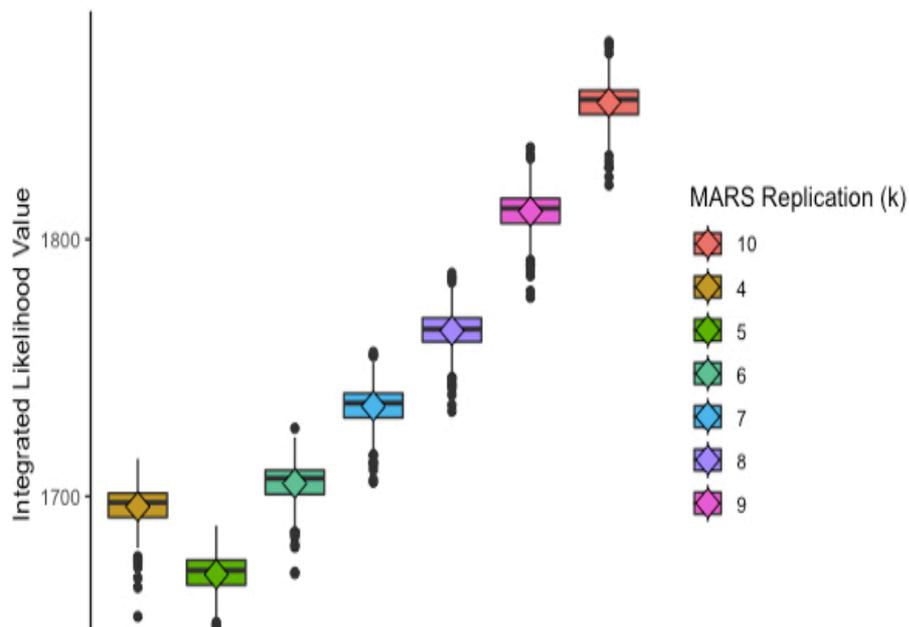
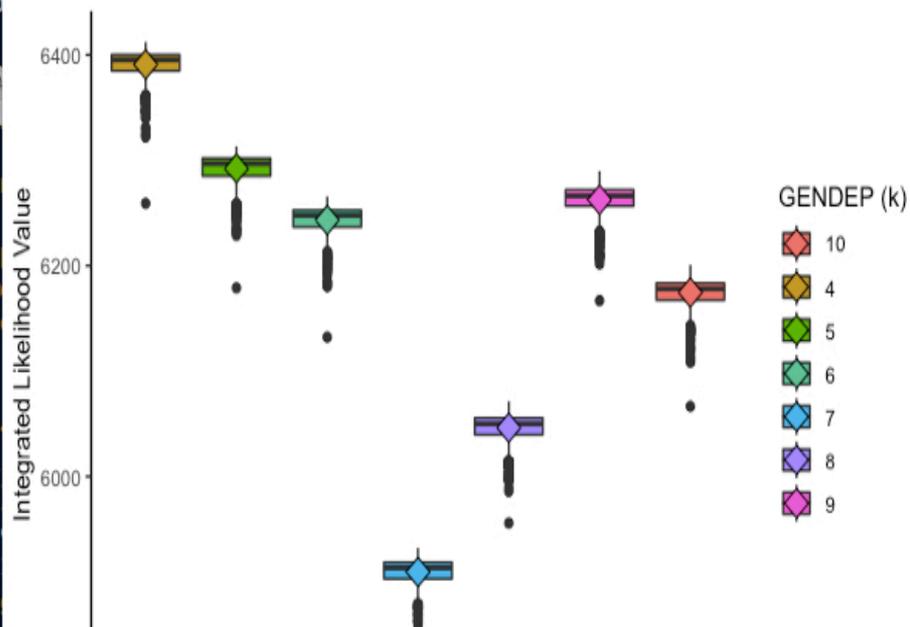
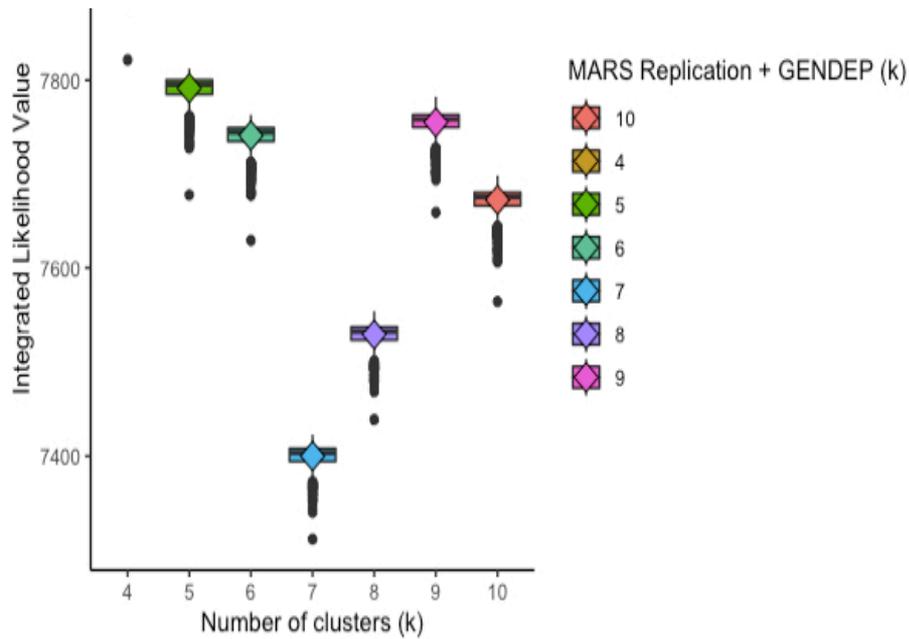
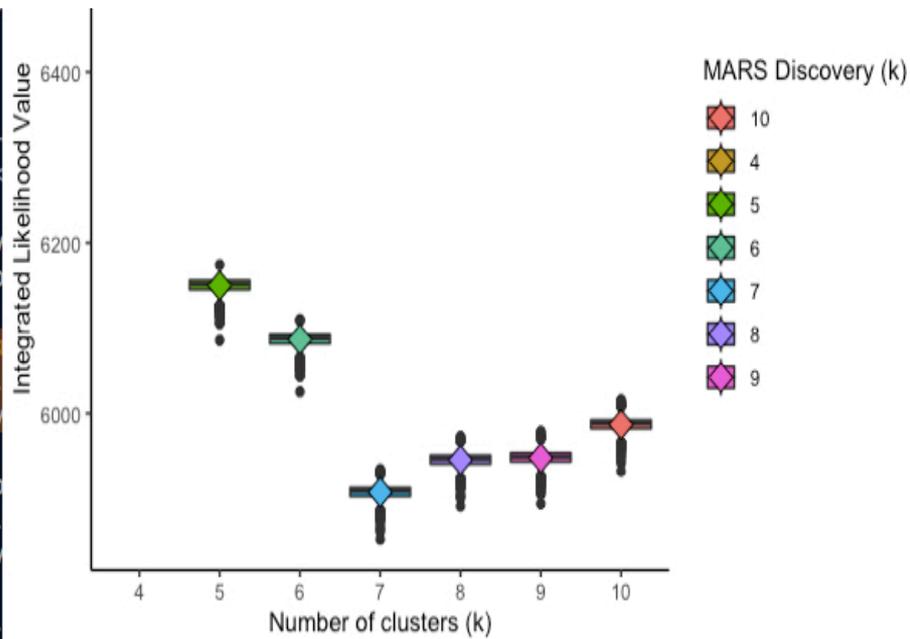


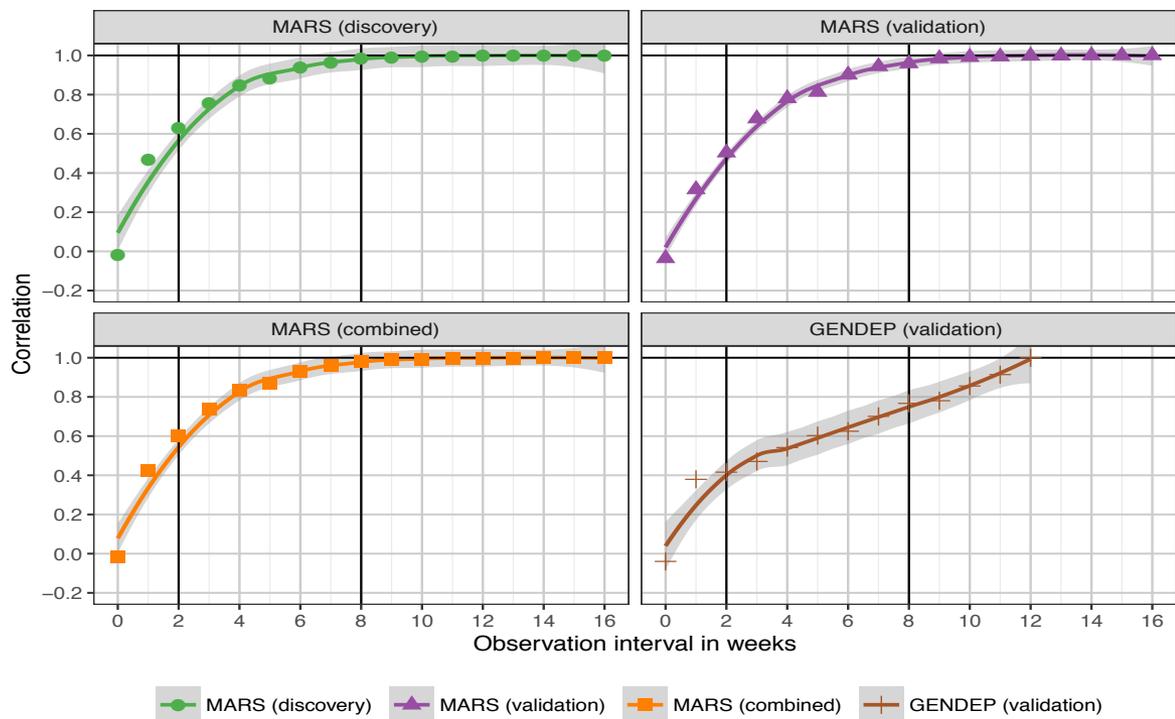
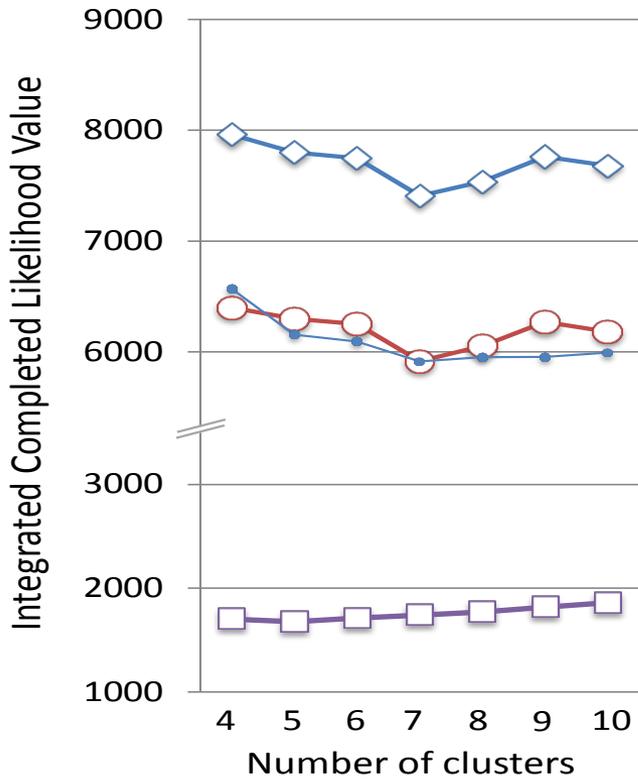
Using Average Slope



Model	Sample N(Dis) =834 N(Rep)=237	R-Squared		Assoc. P-Value		Significance in R -Squared in two different Slopes
		Individual	Average	Individual	Average	
Model 0	Discovery	0.08	0.11	4.17E-12	1.81E-23	0.14
Model 0	Replication	0.06	0.19	3.37E-05	1.41E-12	1.17E-02
Model 1	Discovery	0.08	0.11	1.15E-17	1.22E-23	0.13
Model 1	Replication	0.11	0.20	2.40E-07	3.87E-14	0.06
Model 2	Discovery	0.16	0.21	6.91E-31	1.01E-38	0.04
Model 2	Replication	0.07	0.20	1.72E-05	4.52E-13	0.01
Model 3	Discovery	0.14	0.21	2.09E-29	9.64E-46	0.02
Model 3	Replication	0.11	0.21	1.52E-07	1.05E-14	0.06
Model 4	Discovery	0.25	0.35	8.05E-56	1.37E-85	0.003
Model 4	Replication	0.16	0.27	4.68E-11	2.36E-21	0.04388







- ◇ Combined validation samples (MARS/GENDEP)
- GENDEP replication sample
- MARS discovery sample
- MARS validation sample

Conclusion/Outlook

- **Clustering (LCMM) appears to be workable**
- **Portable to other datasets**
 - Replication
 - GenDep
- **Increases prediction by clinical variables**
 - Post-hoc support for the existence of latent structure?
- **Better performance in replication**
 - Further support for structure
- **GWAS/GW-RF ongoing**