

Introduction

- SAGE-217 [1] is an investigational GABA_A receptor modulator at both synaptic and extrasynaptic receptor subtypes and is currently being studied in patients with psychiatric disorders, including major depressive disorder (MDD) and postpartum depression (PPD). In a pivotal, double-blind, placebo-controlled clinical trial of 89 patients with moderate to severe MDD, administration of SAGE-217 for 14 days was associated with a statistically significant mean reduction in the Hamilton Rating Scale for Depression (HAM-D) 17-Item total score [2]. Recently, in a Phase 3 study in women with PPD, SAGE-217 showed statistically significant improvement in HAM-D score compared to placebo. [3]
- The objective of this analysis was to evaluate alternative models of HAM-D total scores and their predictions of alternative treatment durations, with emphasis on the comparison of prediction accuracy of models that quantify total score directly (TS method) to models that quantify each of the 17 component scores (CS methods) in MDD.

Results

- Alternative methods to model HAM-D scores were evaluated by comparing prediction accuracy for TS and CS methods using multiple model covariates and training/test conditions (see Methods). Compared to the TS method, the CS method demonstrated higher accuracy in all conditions. The RMSE difference of HAM-D total scores between the TS and CS methods was highest for Case 1 (training: D0-7, test: D14), with Model 1 RMSE of 4.67 and 4.16 for the TS method and CS methods, respectively. The RMSE difference was not as large for Case 3 (5-fold cross-validation), with Model 1 RMSE of 3.50 and 3.35 for the TS and CS methods, respectively. Evaluation of the predictions showed that the TS method predicted a few scores that were less than zero (the lower bound of the HAM-D scores), while the CS method predicted all non-negative scores. Detailed evaluation of the model predictions in an example subject showed the benefits of the CS method in providing accurate lower bounds to each component of the HAM-D score, resulting in potentially more accurate predictions.

Figure 1: RMSE comparison showed higher prediction accuracy with the CS method

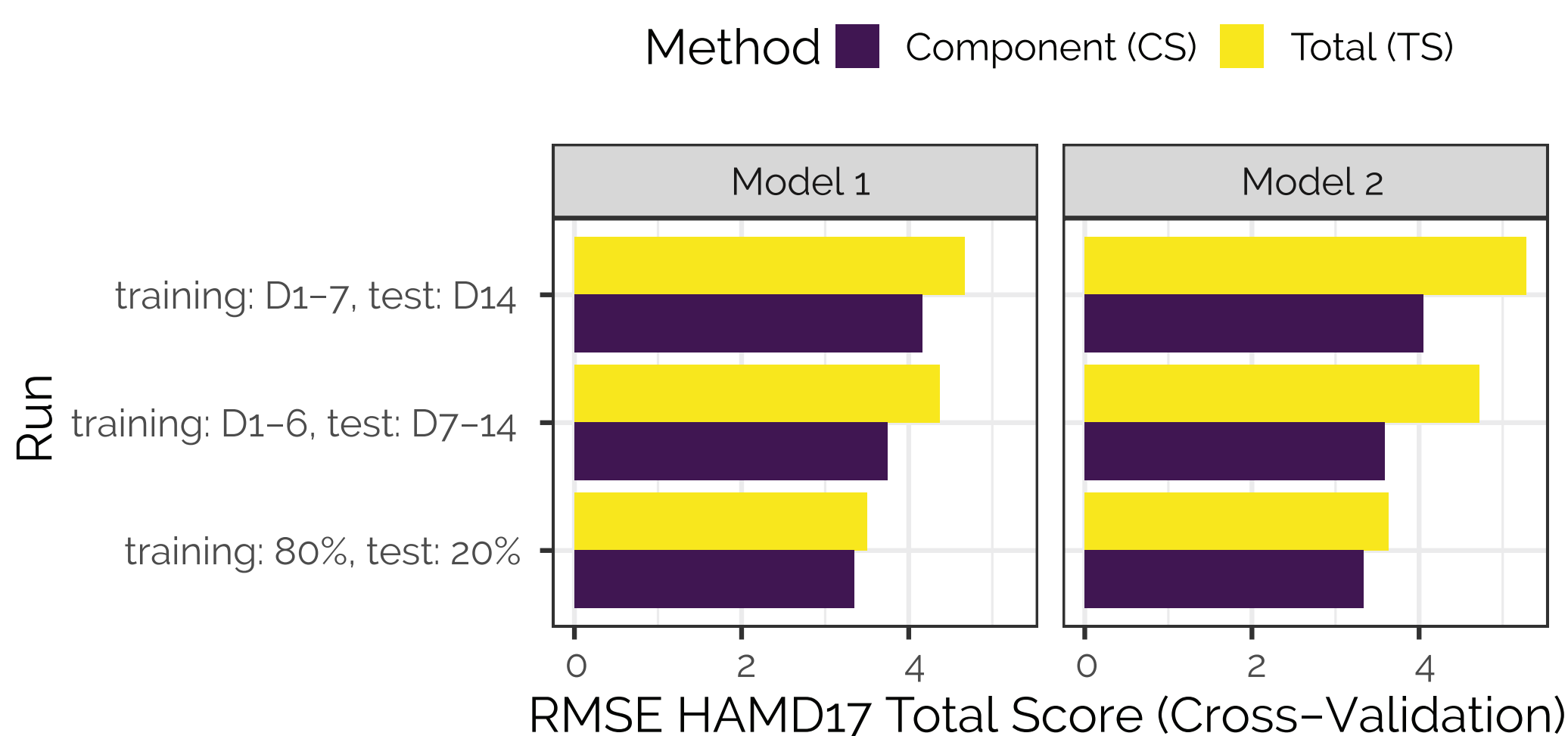
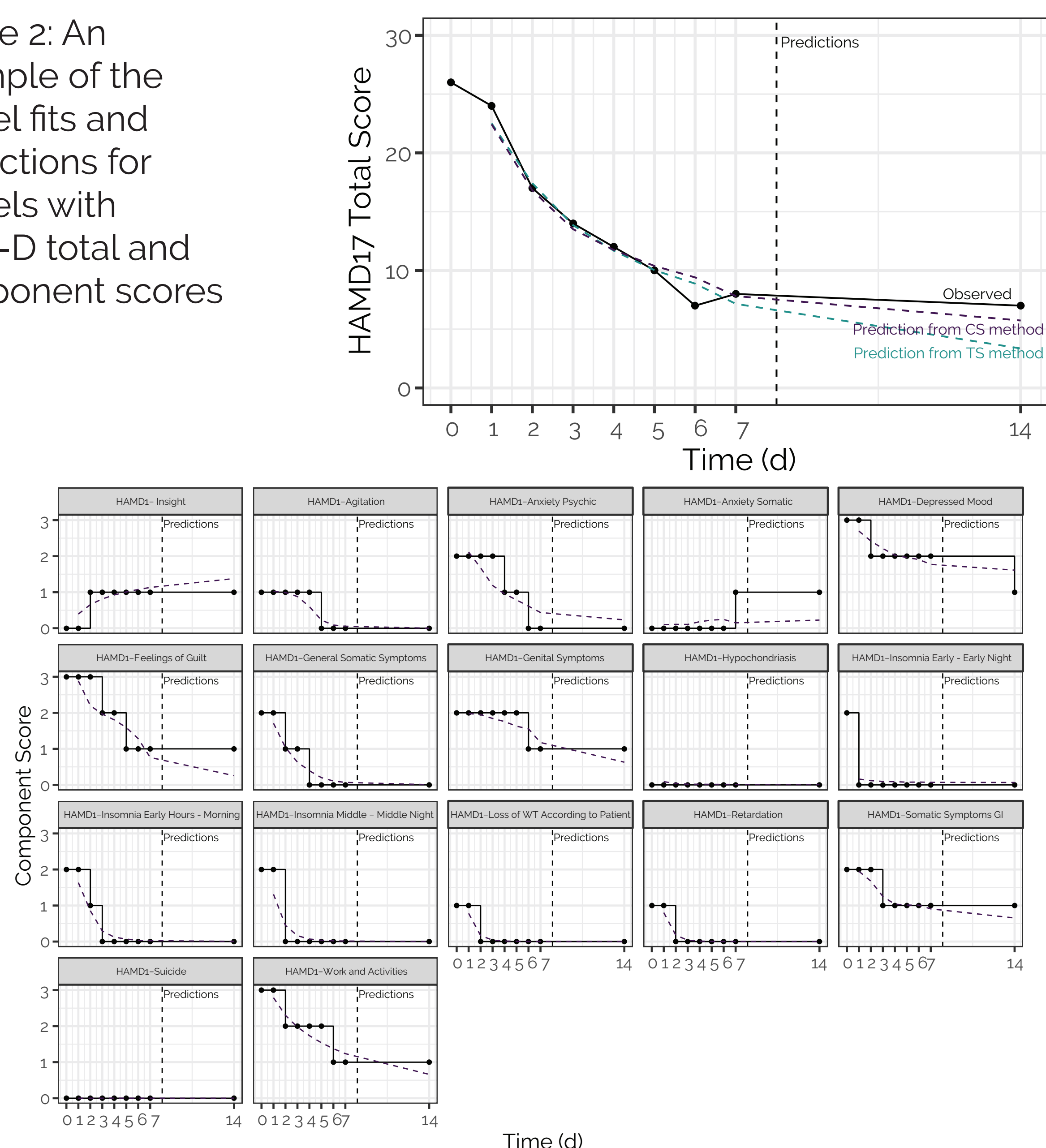


Figure 2: An Example of the model fits and predictions for models with HAM-D total and component scores



Methods

- Baseline (Day 0 [D0]) and on-treatment HAM-D assessments on Days 1-7, and 14 from 89 MDD patients (44 placebo, 45 SAGE-217) were used to evaluate alternative methods. Two mixed-effect models were tested: Model 1 (primary): HAM-D~cumulative concentration (log-transformed), and Model 2 (secondary): HAM-D~time (log-transformed). Regression estimates were corrected by the placebo effects in each timepoint. The use of log-transformed cumulative concentration was based on the apparent linear relationship between total score estimate (with discrete time) and log-scale cumulative PK (Figure S1). In the TS method, HAM-D total scores were modeled as continuous variables. In the CS method, HAM-D component scores were modeled as ordered categorical variables. The concordance of the observed and predicted total score from the CS method is given in Figure S2.
- Prediction accuracy was assessed by comparing root mean square error (RMSE) of observed and predicted HAM-D total score in the test set, from the estimates of the training set. Three cases of training/test sets were evaluated: 1) training: D0-7, test: D14; 2) training: D0-6, test: D7 & D14; and 3) training: 80%, test: 20% (5-fold cross-validation). Comparisons of the TS and CS methods were performed by evaluating model performance from the same training/test sets.

Figure S1: Total score is log-linearly related to cumulative PK

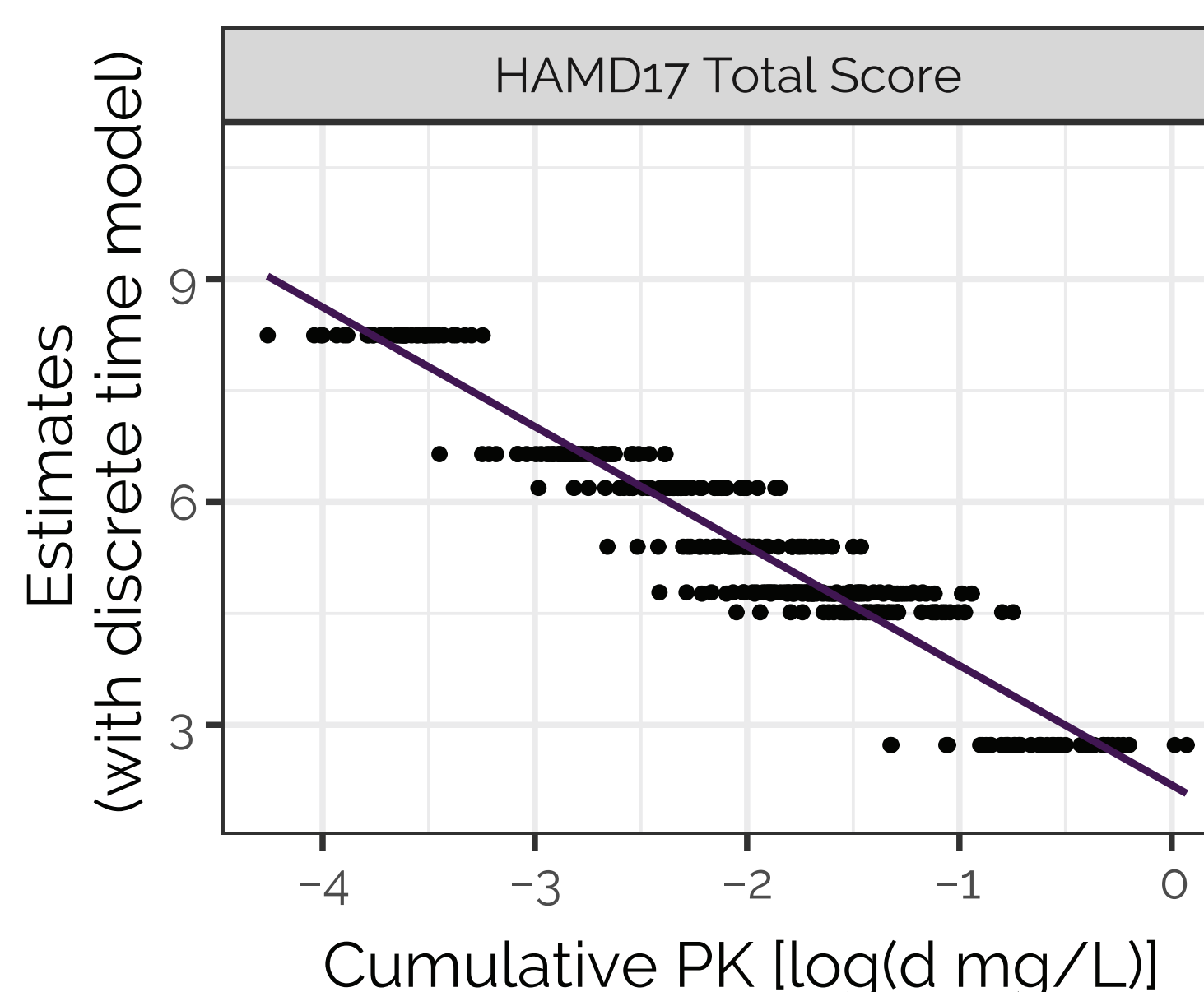
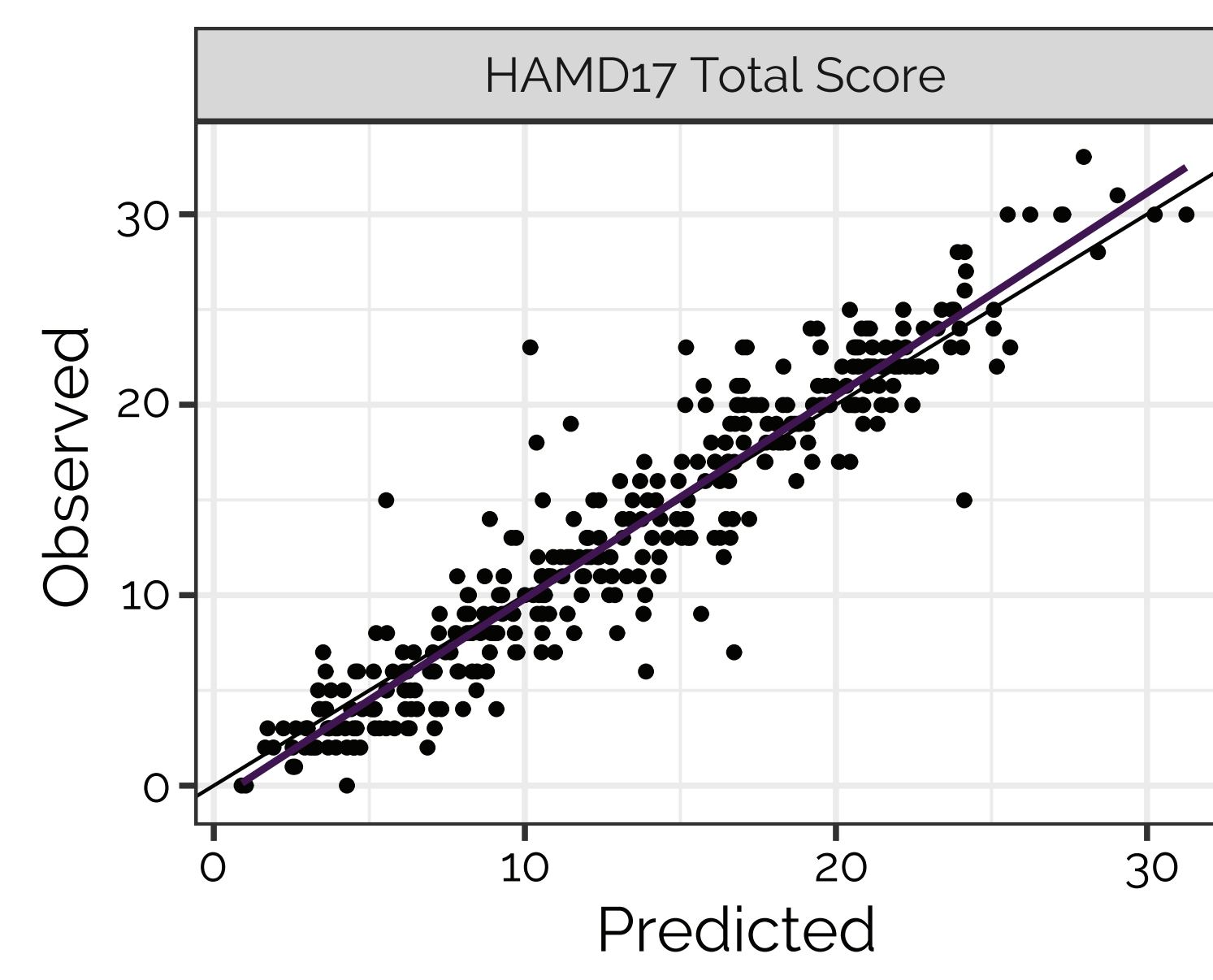


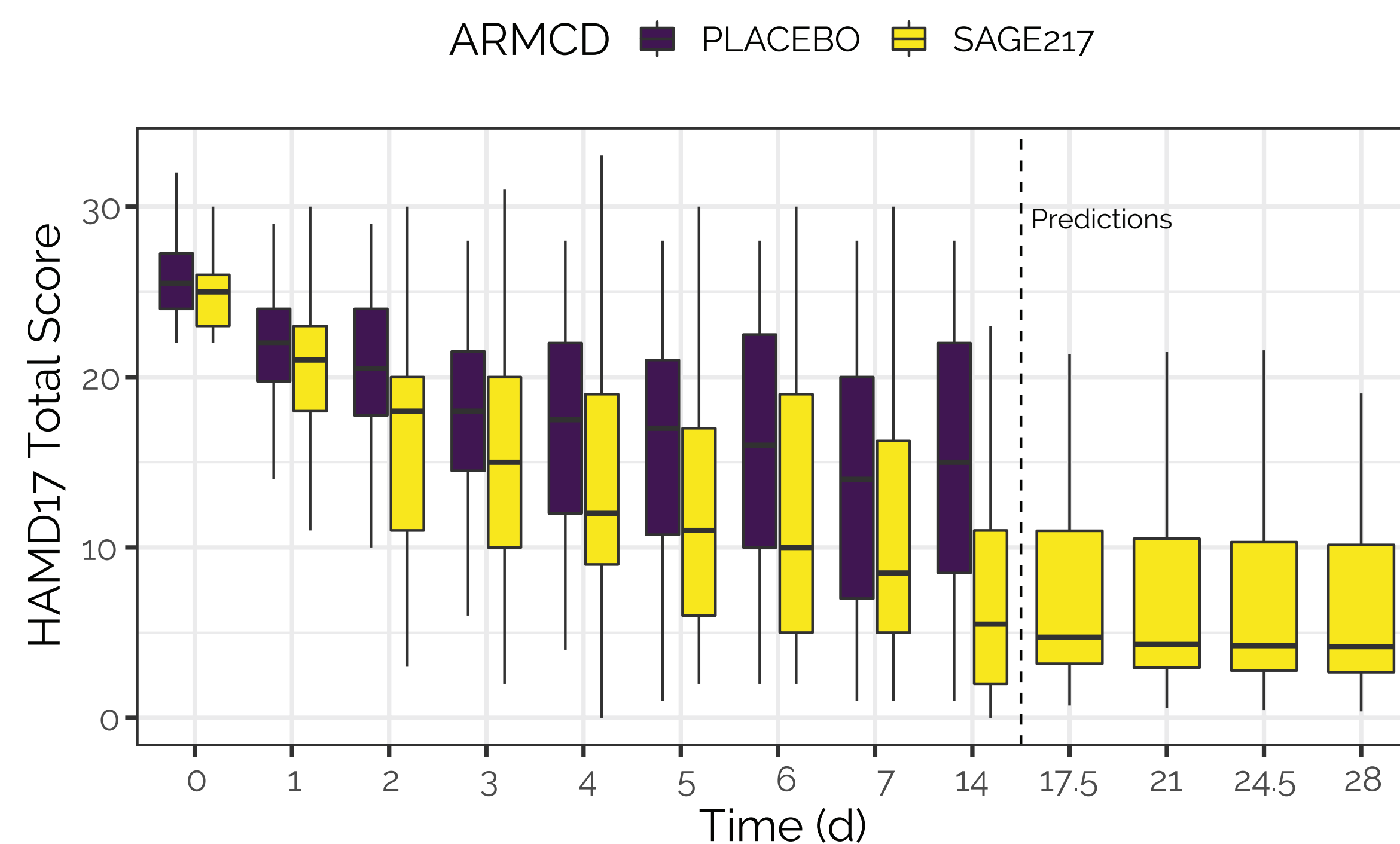
Figure S2: Observed and Predicted total score from CS methods



Predictions

- The CS method was used to make predictions of HAM-D total scores for alternative duration of SAGE-217 treatment beyond 14 days. The results indicated that the optimal treatment duration in MDD was 14 days; prolonged durations of treatment beyond 14 days were not predicted to improve HAM-D total scores. The TS method may have incorrectly predicted some improvements with longer durations.

Figure 3: Prediction of HAM-D Total Score by Duration of Treatment



Conclusions

- HAM-D total scores in patients with MDD were predicted more accurately by quantifying each HAM-D component score than compared to quantifying HAM-D total scores directly. The results may permit more accurate study planning during clinical development.

REFERENCES

1. M 1 Martinez Botella G, et al. J Med Chem. 2017 Sep 28;60(18):7810-7819. doi: 10.1021/acs.jmedchem.7b00846. Epub 2017 Aug 17. 2 Gunduz-Brace, G., et al. 2018 ENCP Congress, Barcelona, Spain. <https://www.ecnp.eu/presentationpdfs/73/P032.pdf> 3. Sage Therapeutics press release, Jan 7 2019

DISCLOSURES

BBS, LR, and SAJ are employees of Sage Therapeutics, Inc., with stock/stock options.

SUPPORT

These studies were funded by Sage Therapeutics, Inc. Editorial support in the editing, design, and production of this poster was provided by Boston Strategic Partners, Inc. and supported by Sage Therapeutics, Inc.

SAGE THERAPEUTICS and the SAGE THERAPEUTICS logos are trademarks of Sage Therapeutics, Inc.

Presented at the 2019 Annual Meeting of the Population Approach Group Europe, Stockholm, Sweden June 11-14, 2019

MRC-217-00422