UPPSALA UNIVERSITET

The LASSO – A Novel Method for Predictive Covariate Model Building in Nonlinear Mixed Effects Models

Jakob Ribbing, Joakim Nyberg, Ola Caster and E. Niclas Jonsson

Division of Pharmacokinetics and Drug Therapy
Department of Pharmaceutical Biosciences
Uppsala University

Introduction - Covariate Selection in Nonlinear Mixed Effects Models

UPPSALA UNIVERSITET

- Stepwise-Covariate Modelling (SCM)
- Some of the problems with SCM:

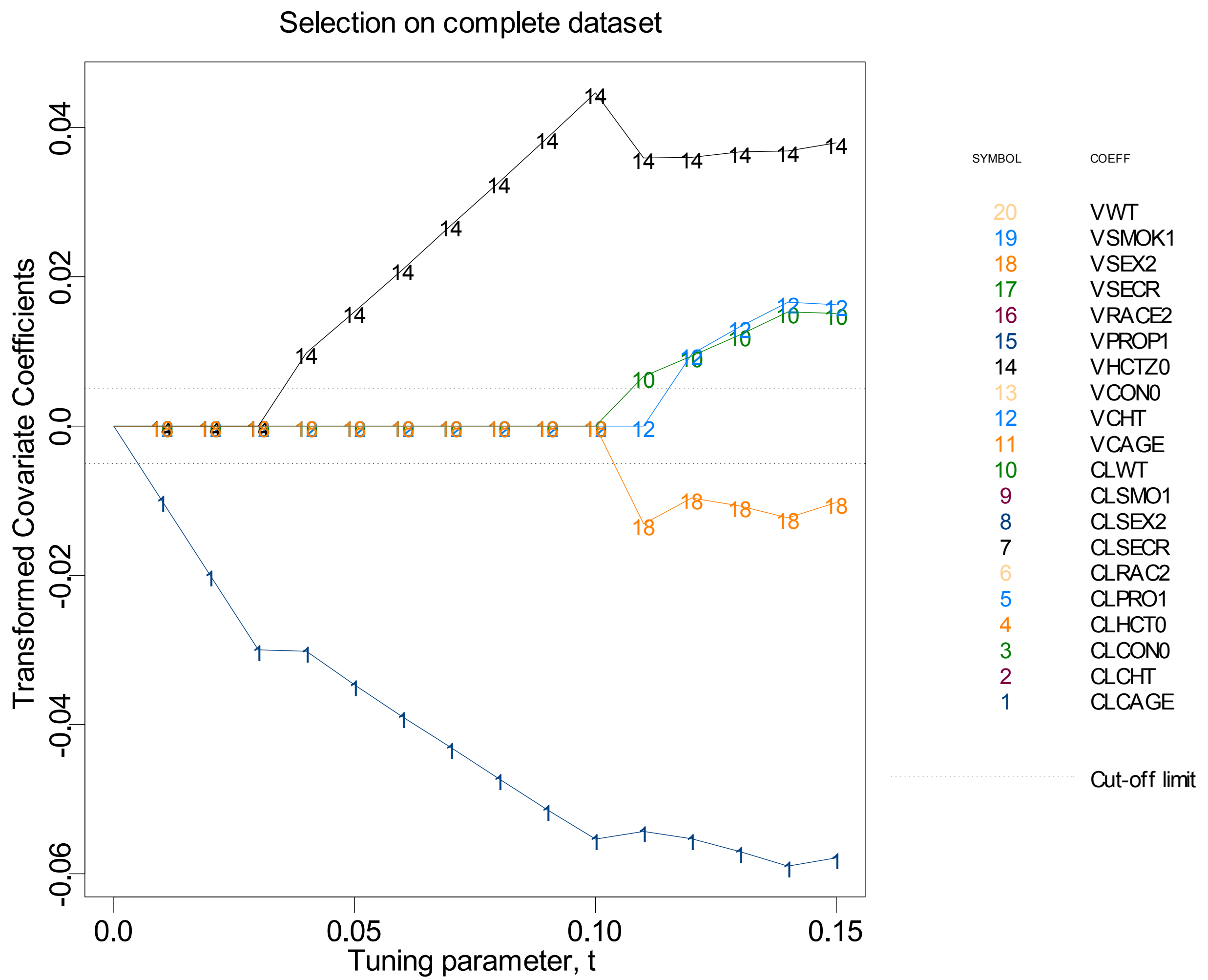
1. Border-line significant covariate effects either discarded or included
   - ➢ The LASSO would shrink these covariate effects but may keep them in the model
2. User must specify p-value for selection
   - ➢ The LASSO uses cross-validation
3. Long computer-run-times
   - ➢ The LASSO may be faster

UPPSALA UNIVERSITET

Theory - LASSO for “Least Absolute Shrinkage and Selection Operator”

- Covariate transformation
  - centred around zero
  - normalised to between-individual standard deviation (unless time-varying covariates)
- Covariate-coefficient magnitude on same scale
- Estimating the lasso model:
  - estimating full covariate model with restriction
    - absolute sum of covariate coefficients ≤ t
  - t (tuning parameter), determines Model Size

Theory - Illustration of the LASSO-Estimates over t



Objectives

To implement the LASSO for covariate selection within NONMEM and to compare this method to the commonly-used SCM

Method – Implementation of the LASSO

- Implemented as a fully automated tool using Perl-speaks-NONMEM (PsN)

- Optimal t estimated using cross-validation
  - Cross validation similar to data splitting but uses data more efficiently
  - Five-fold cross validation on NONMEM objective function value (OFV)

Method – Creation of Analysis Datasets

- Analysis datasets generated by sampling subjects (with replacement) from a PK dataset containing 721 subjects
  - 40, 60, 120 or 180 subjects in each analysis dataset
  - 100 replicate dataset of each size

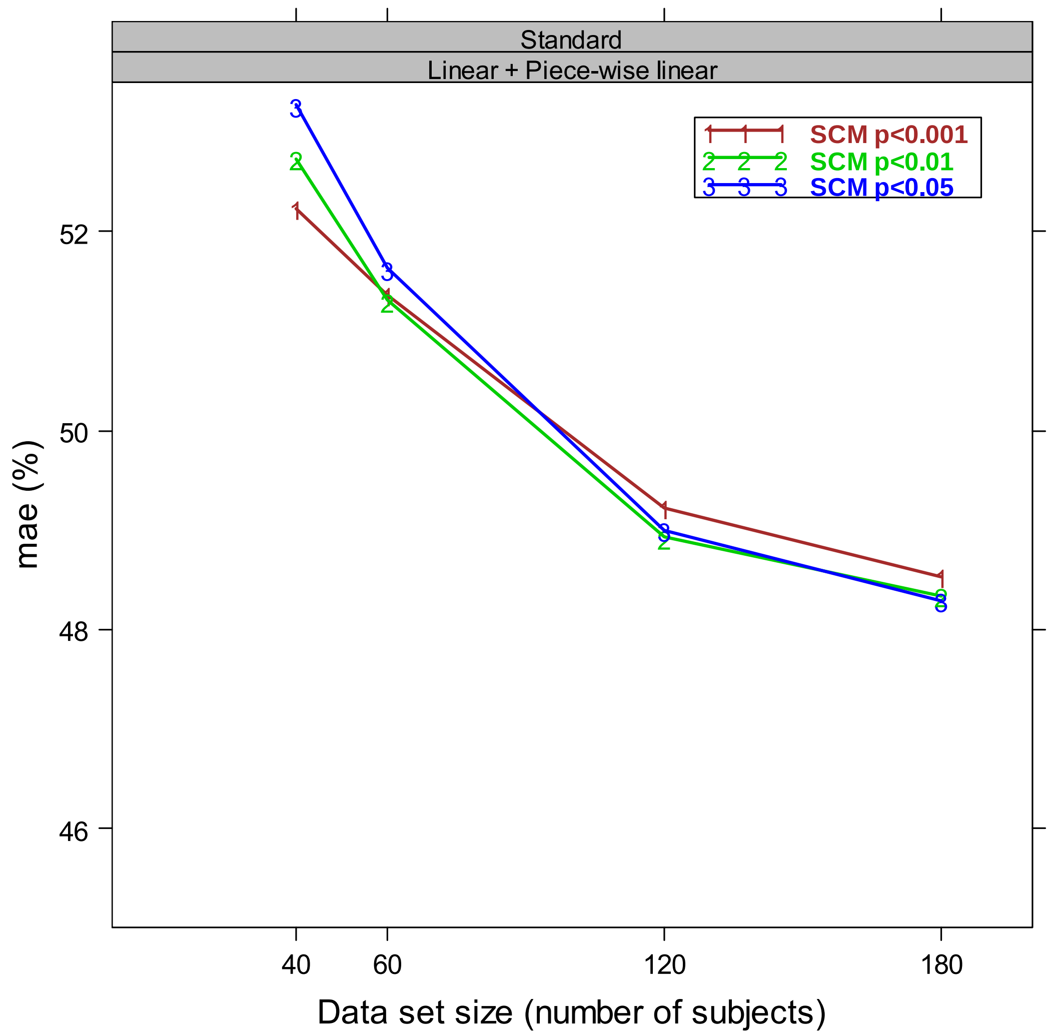
Method – Creation of Validation Datasets

- For each analysis dataset a validation dataset was created comprising all subjects among the 721 that were not in the corresponding analysis dataset

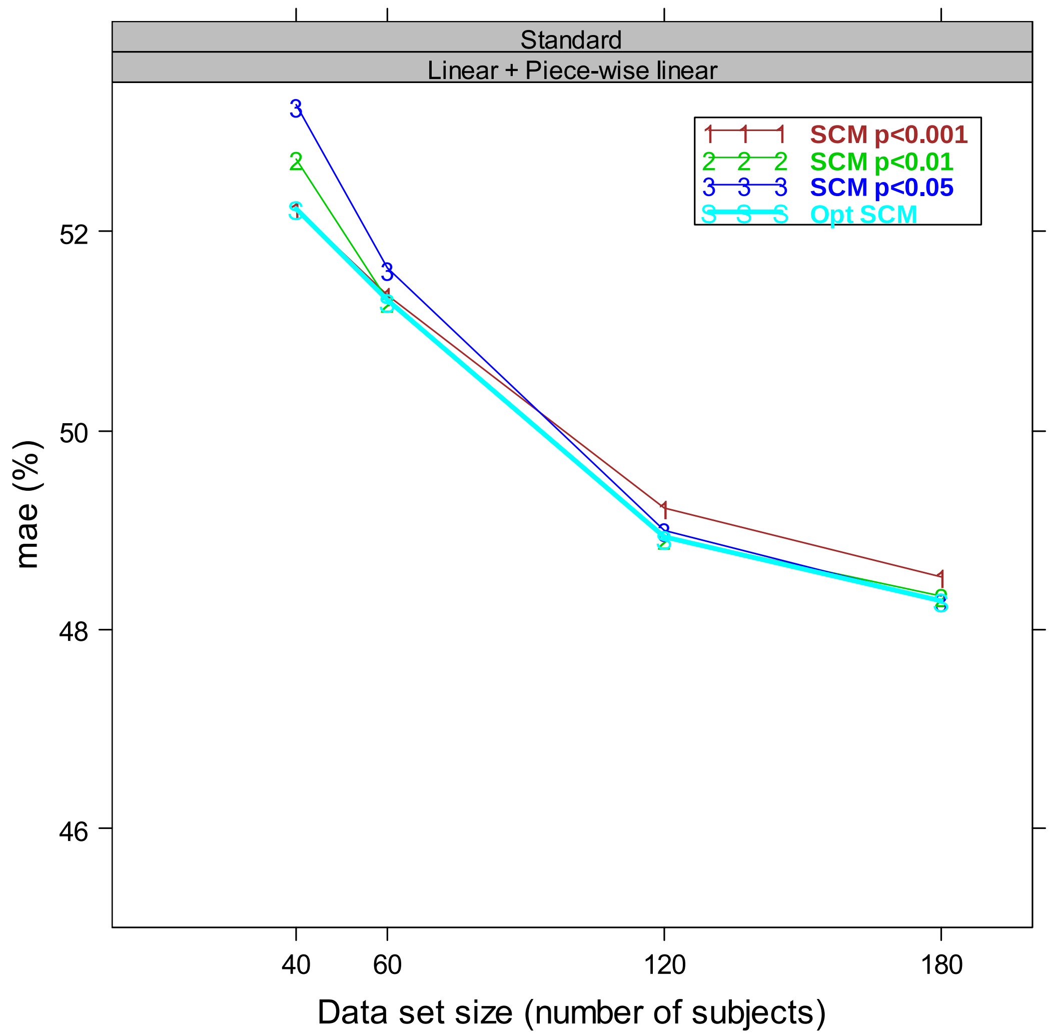
- To compare models produced by SCM and LASSO, prediction error evaluated on observations in validation dataset:

    - mae = average($|obs_n - pred_n| / obs_n$)$\cdot$100%

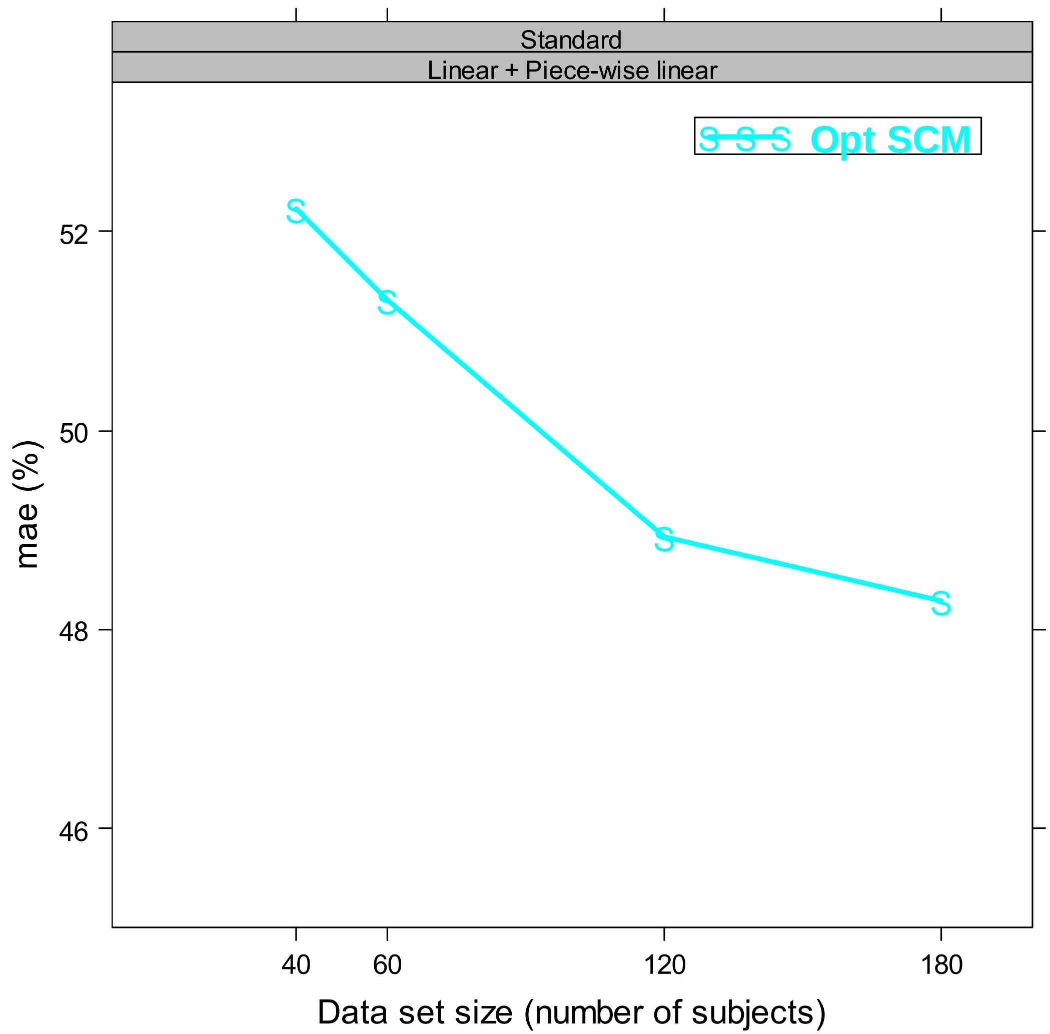
Results – Prediction Error for the SCM with Different p-values



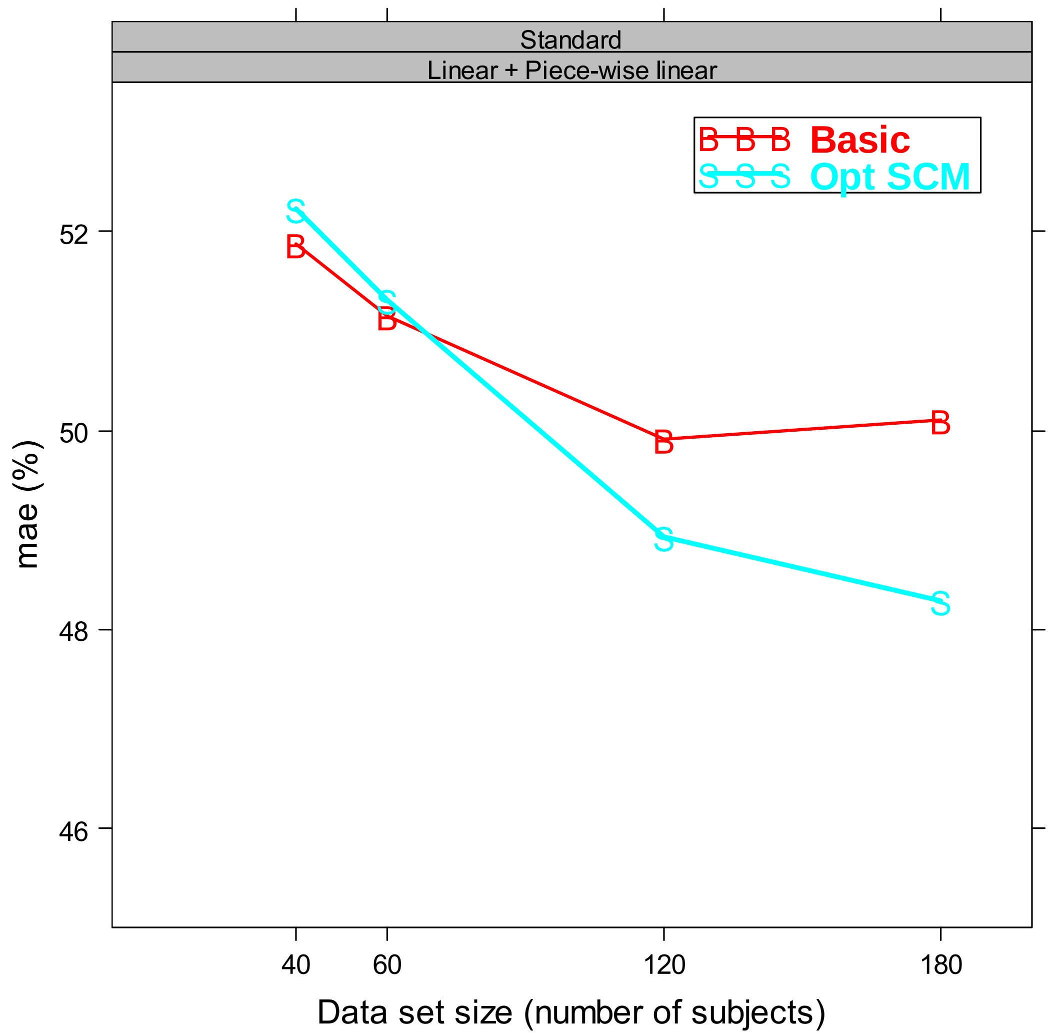
Results – Prediction Error for the SCM with Different P-Values



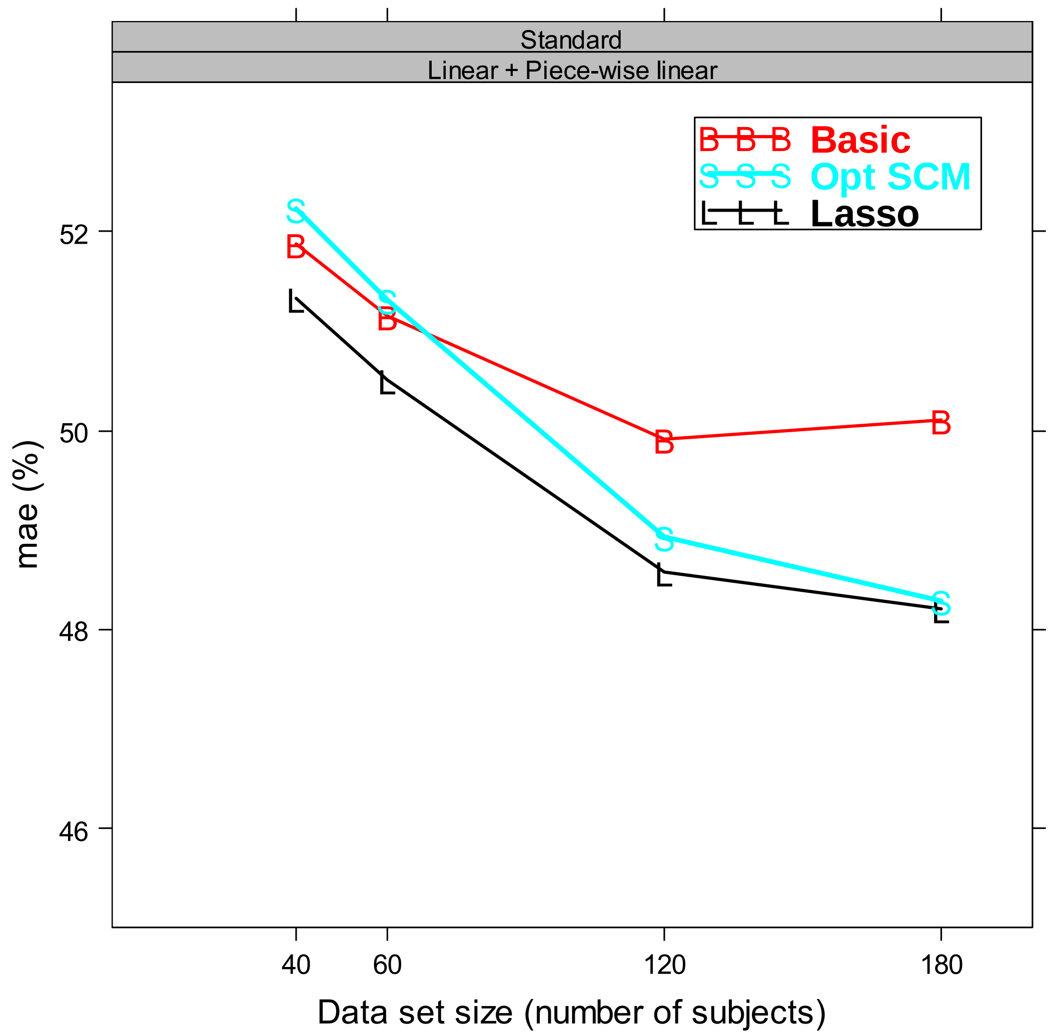
Results – Prediction Error for SCM



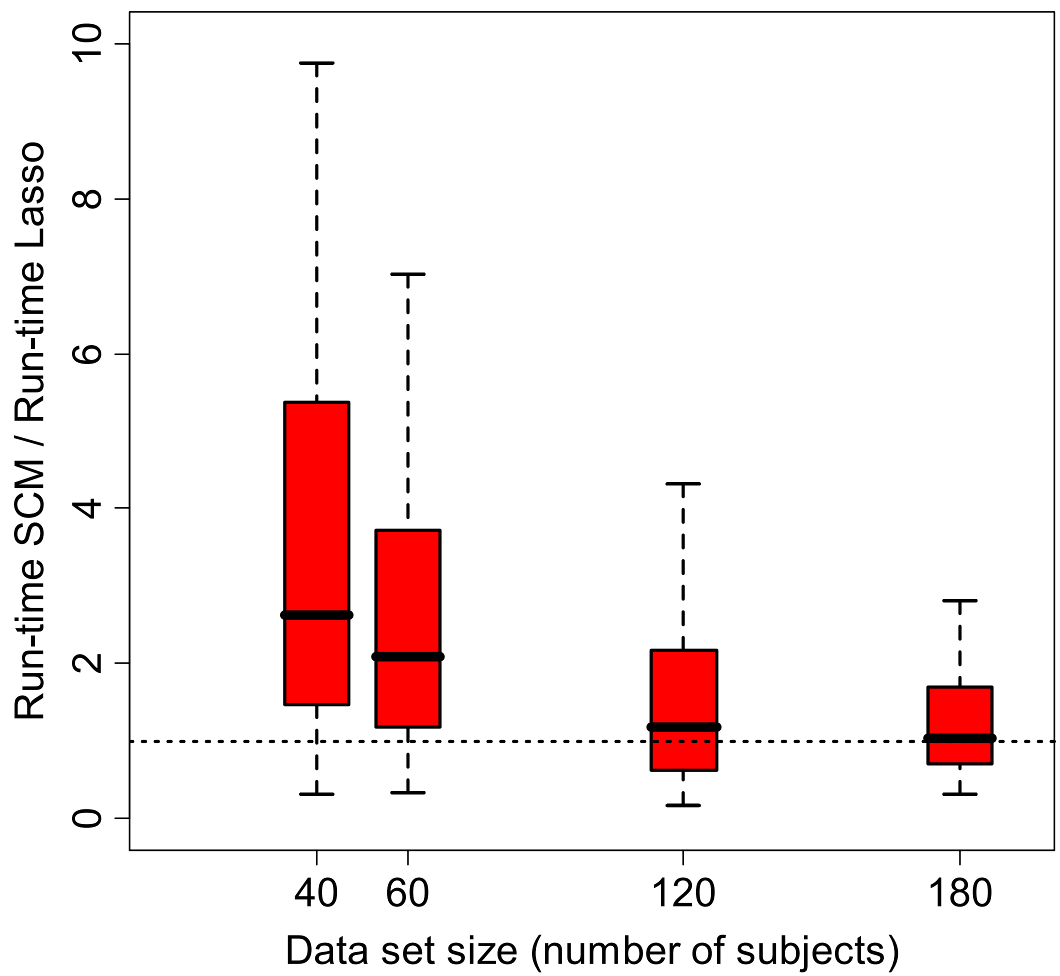
Results – Prediction Error for SCM and Starting Model



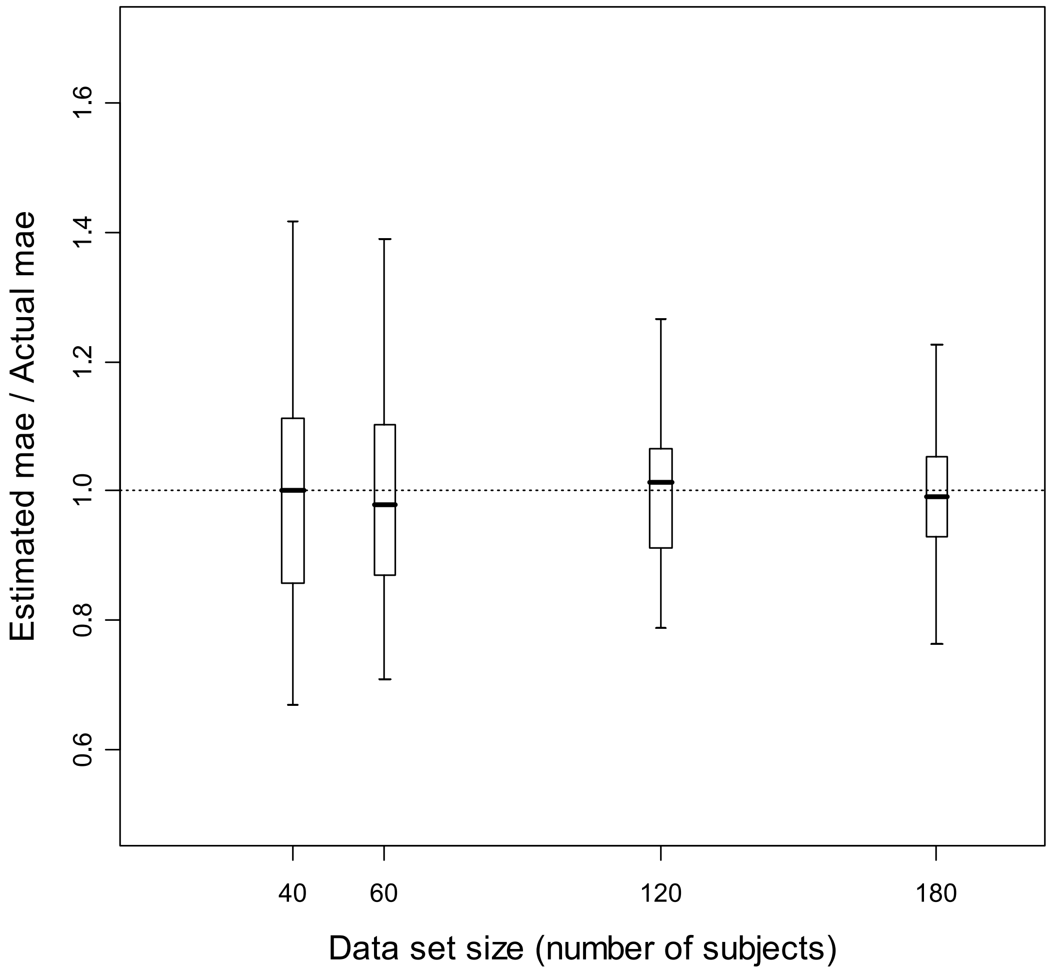
Results – Prediction Error for SCM, LASSO & Starting Model



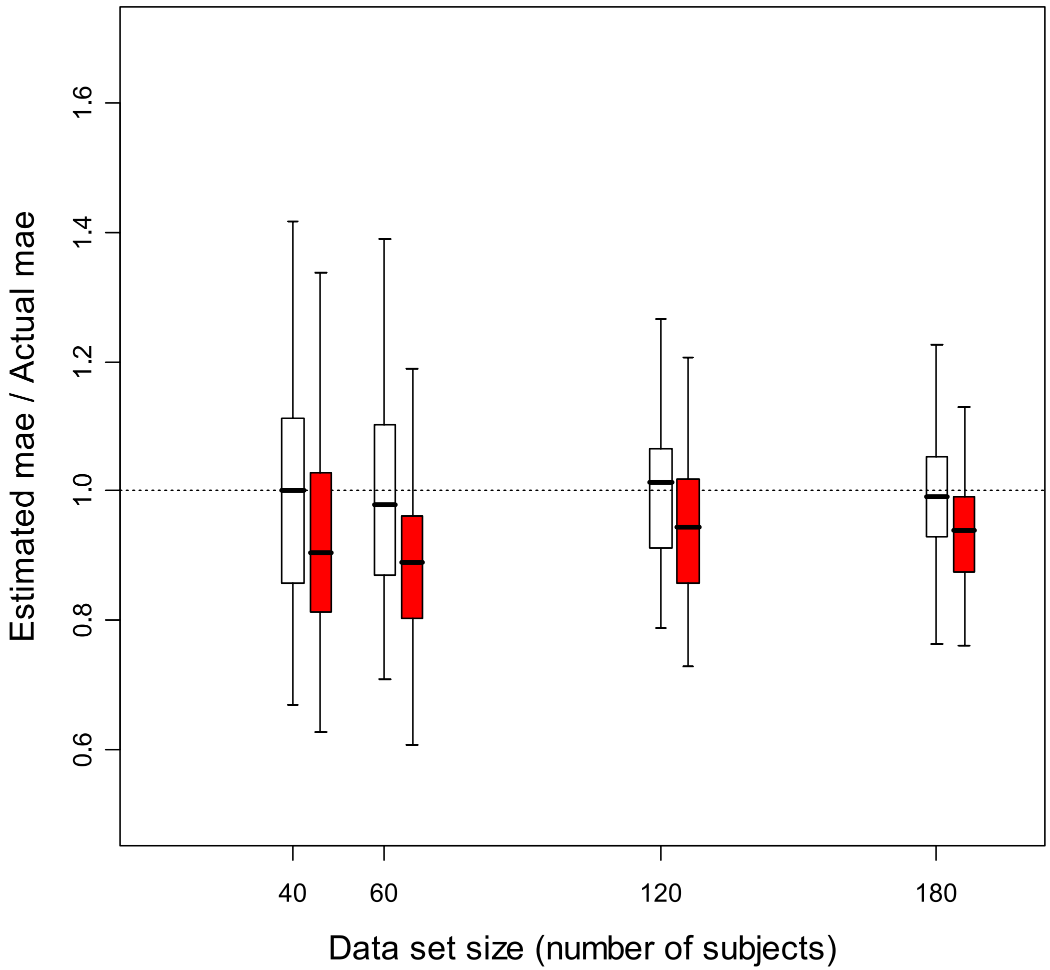
Results – Computer Run-Time



Results – LASSO Provides Unbiased Estimate of Prediction Error



Results – SCM Provides No Accurate Estimate of Prediction Error



Discussion and Conclusions – Drawbacks of the LASSO

- May produce a more complex model
- Cross-validation difficult on unstable model
  - Estimable on 80% of the original data
- Little experience of this method in pop PK/PD

Discussion and Conclusions – Advantages of the LASSO

The LASSO is preferable for small datasets

- Better predictive performance
  - Also for small subpopulations in large datasets!
- Shorter run-time if many covariate relations

- No need to specify a p-value for selection

- Provides estimate of prediction error
  - Covariate-model selection taken into account
  - External validation of covariate model!

Take-Home Message

On small datasets use the LASSO rather than the SCM

UPPSALA UNIVERSITET

Acknowledgements

- AstraZeneca and Per-Henrik Zingmark
  - Data
- Lars Lindbom and Pontus Pihlgren
  - Cluster and PsN
- Mats Karlsson and Nicholas Holford
  - Valuable discussion and suggestions