

High AUROC, high instability

Title: The kitchen-sink approach may undermine individual-level decisions

Background: Within OHDSI, a “kitchen sink” modeling approach is commonly recommended, in which all candidate predictors are included, and the data is allowed to determine their relative importance for predicting a given outcome. This strategy is often justified by observed improvements in discrimination, such as higher AUROC. However, does the kitchen sink approach also reduce model uncertainty on individual level?

Methods

1 Data sources:

1. Merative MarketScan® MultiState Medicaid Database (MDCD).
2. The Integrated Primary Care Information (IPCI) database.

2 Prediction tasks:

1. Predicting colorectal cancer
2. Predicting dementia
3. Predicting lung cancer

3 Candidate predictors:

1. Age + sex
2. Constrained predictor set (n =51)
3. Kitchen sink approach (unrestricted)

4 Model choice:

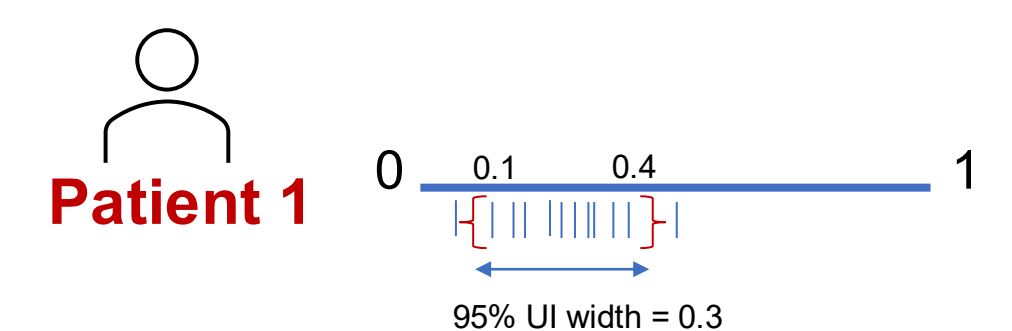
1. LASSO

5 Model development:

1. Learning curves using 350-5,000 outcome events as sample sizes
2. Bootstrap all sample sizes 200 times to estimate individual level instability

6 Metrics used:

1. AUROC
2. 95% Uncertainty interval (UI) width:
per risk group (quantiles: 1: low risk,
5: high risk)
3. Probability of misclassification



Results

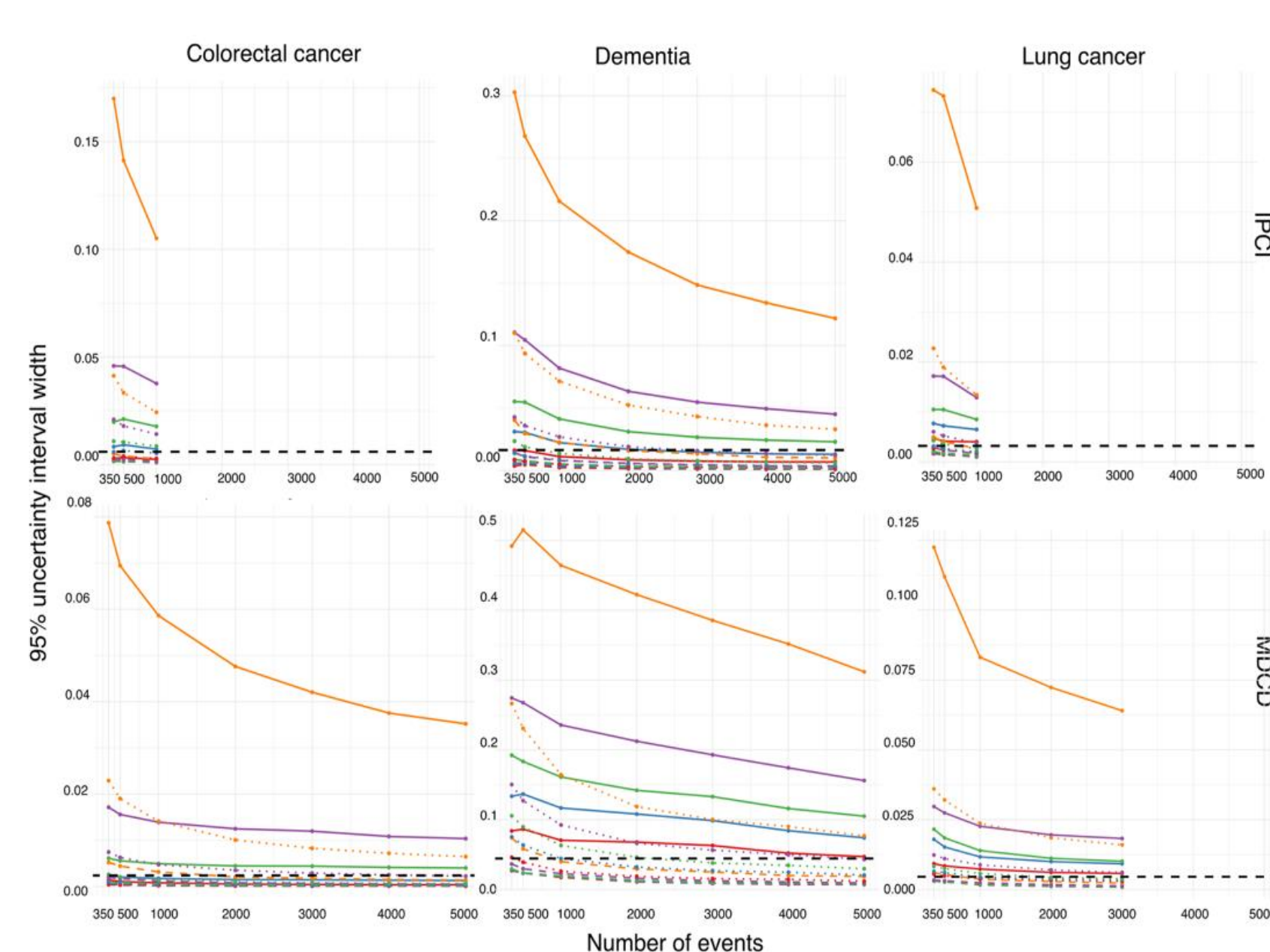


Figure 1 95% UI width per risk quantile on the probability scale

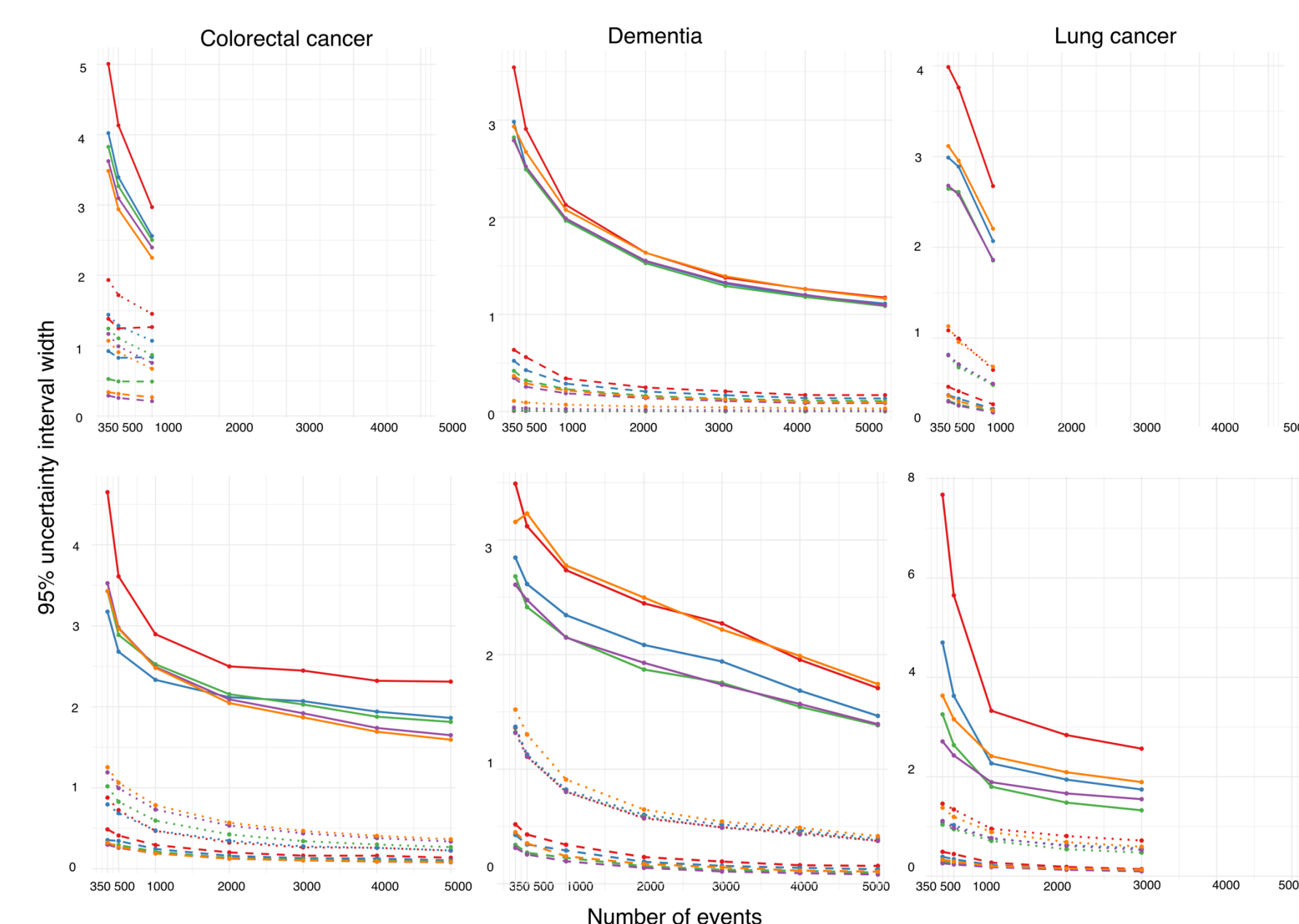


Figure 2 95% UI width per risk quantile on the logit scale

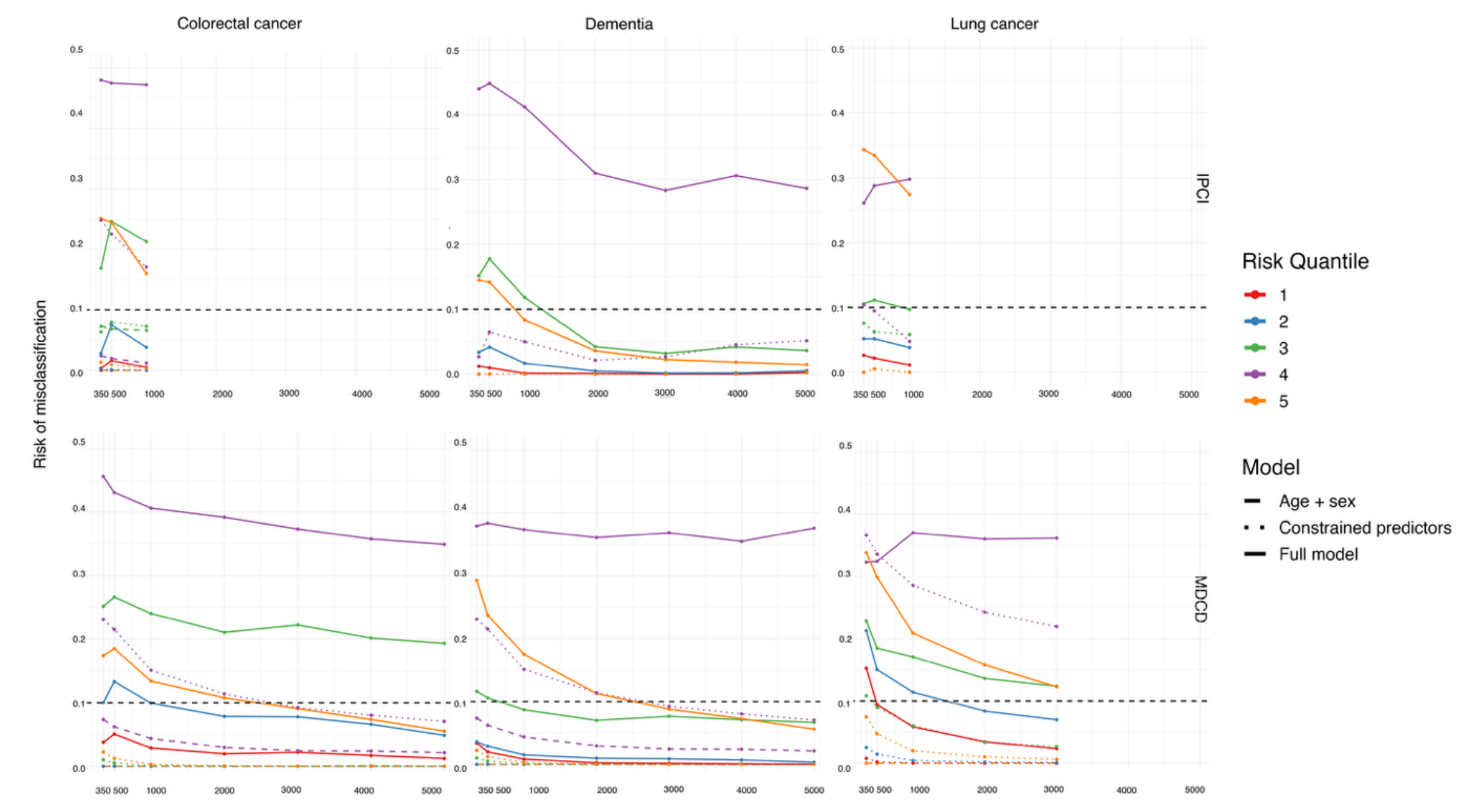


Figure 3 Probability of misclassification. Defined as proportion of bootstraps leading to a different classification than the reference model.

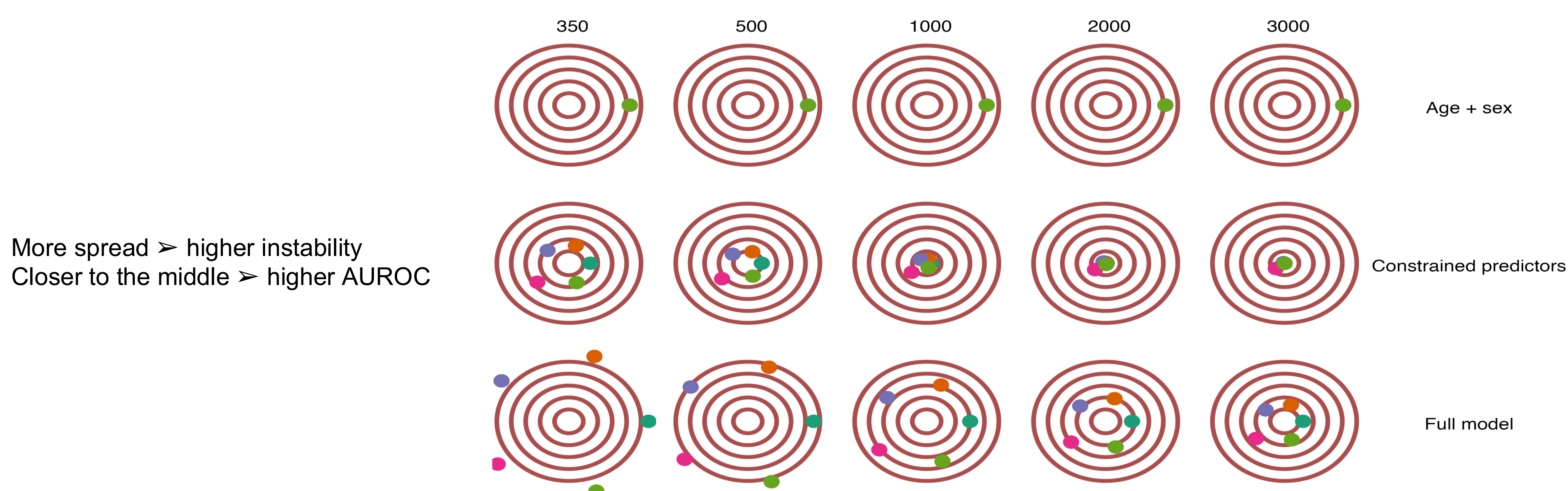


Figure 4 Trade off AUROC vs instability

Conclusion: The kitchen sink approach increases overfitting and individual instability in prediction models