

## Incorporating published univariable associations in diagnostic and prognostic modeling

### INTRODUCTION

Diagnostic and prognostic literature is overwhelmed with studies reporting univariable predictor-outcome associations. Currently, methods to incorporate such information in the construction of a prediction model are underdeveloped and unfamiliar to many researchers.

### METHODS

This article aims to improve upon an adaptation method originally proposed by Greenland (1987) and Steyerberg (2000) to incorporate previously published univariable associations in the construction of a novel multivariable prediction model.

### GREENLAND/STEYERBERG ADAPTATION METHOD

Univariate (u) → adjusted (m) RR [2,3]

I = IPD, L = Literature + IPD

$$\beta_{m|L} = \beta_{u|L} + (\beta_{m|I} - \beta_{u|I})$$

$$\text{var}(\beta_{m|L}) = \text{var}(\beta_{u|L}) + \text{var}(\beta_{m|I}) - \text{var}(\beta_{u|I})$$

### IMPROVED ADAPTATION METHOD

► Unbiased variance component

$$\text{var}(\beta_{m|L}) = \text{var}(\beta_{u|L}) + \text{var}(\beta_{m|I}) + \text{var}(\beta_{u|I}) - 2\text{cov}(\beta_{m|I}, \beta_{u|I})$$

► Distributional

$$\beta_{u|L} \sim \mathcal{N}(\mu_{u|L}, \sigma_{u|L}^2), \beta_{m|I} \sim \mathcal{N}(\mu_{m|I}, \sigma_{m|I}^2), \beta_{u|I} \sim \mathcal{N}(\mu_{u|I}, \sigma_{u|I}^2)$$

► Robust estimation [1]

$$\mu_{m|I} \sim \text{Cauchy}(0, 2.5), \mu_{u|I} \sim \text{Cauchy}(0, 2.5)$$

### SIMULATION STUDY

- Reference model with 2 predictors for generating data with  $x_1, x_2 \sim \mathcal{N}(0, 1)$  and  $r(x_1, x_2) = 0$
- Individual Patient Data ( $n_{IPD} = 100 \rightarrow 1000$ )
- 4 heterogeneous literature studies ( $n_j = 500$ )

### APPLICATION

- Prediction of peri-operative mortality (in-hospital or within 30 days) after elective abdominal aortic aneurysm surgery
- Individual patient data: 238 patients (including 18 deaths)
- Univariable literature data: 15 studies with 15,821 patients (including 1,153 deaths)

### CONCLUSION

- Adaptation methods outperform established approaches ignoring prior evidence
- Adaptation methods are especially worthwhile when relatively limited individual patient data are available

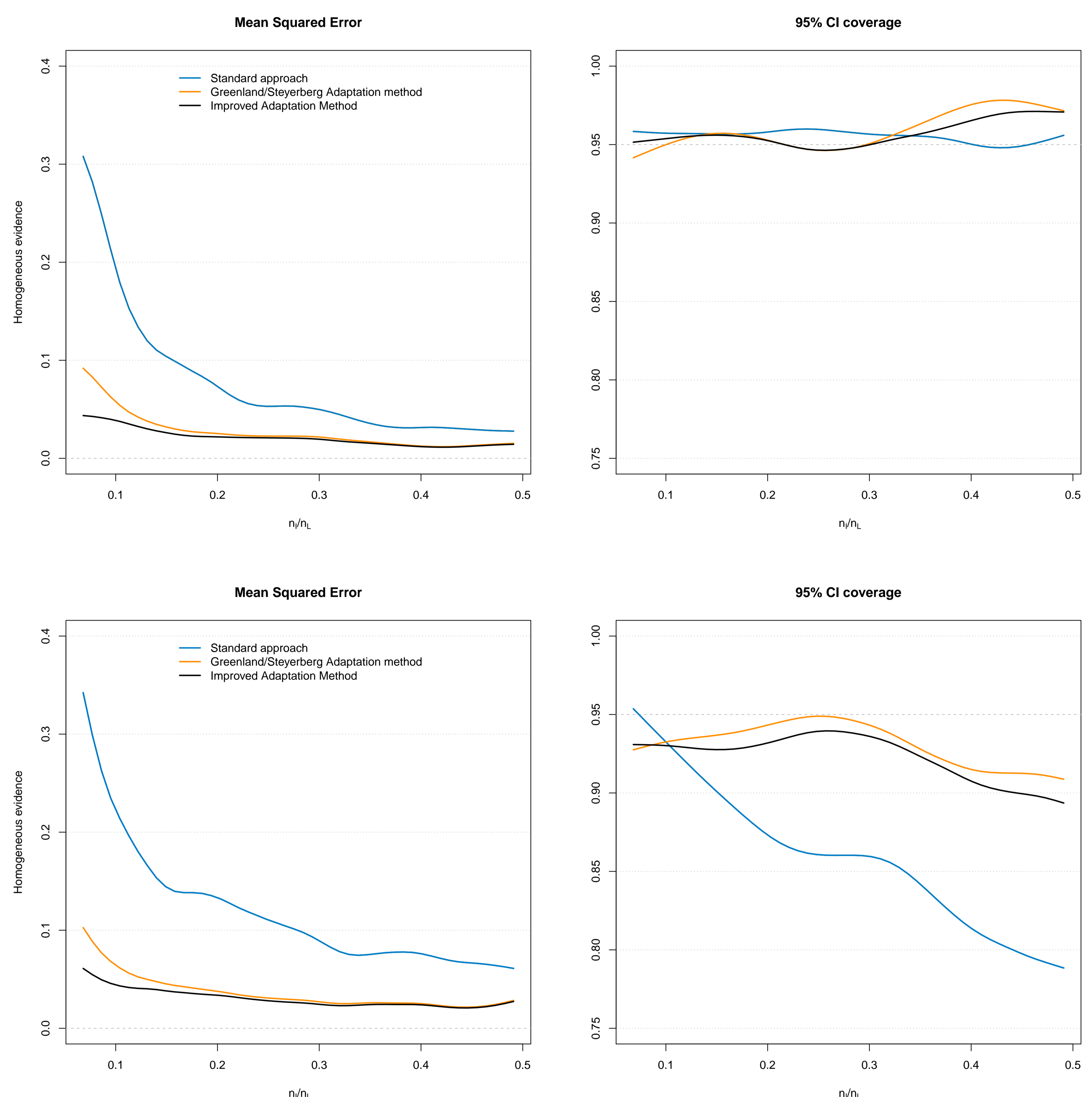
### CONTACT



Thomas Debray, MSc  
Julius Center for Health Sciences and Primary Care  
University Medical Center Utrecht  
Tel: +31 (0) 88 75 68640  
Email: T.Debray@umcutrecht.nl

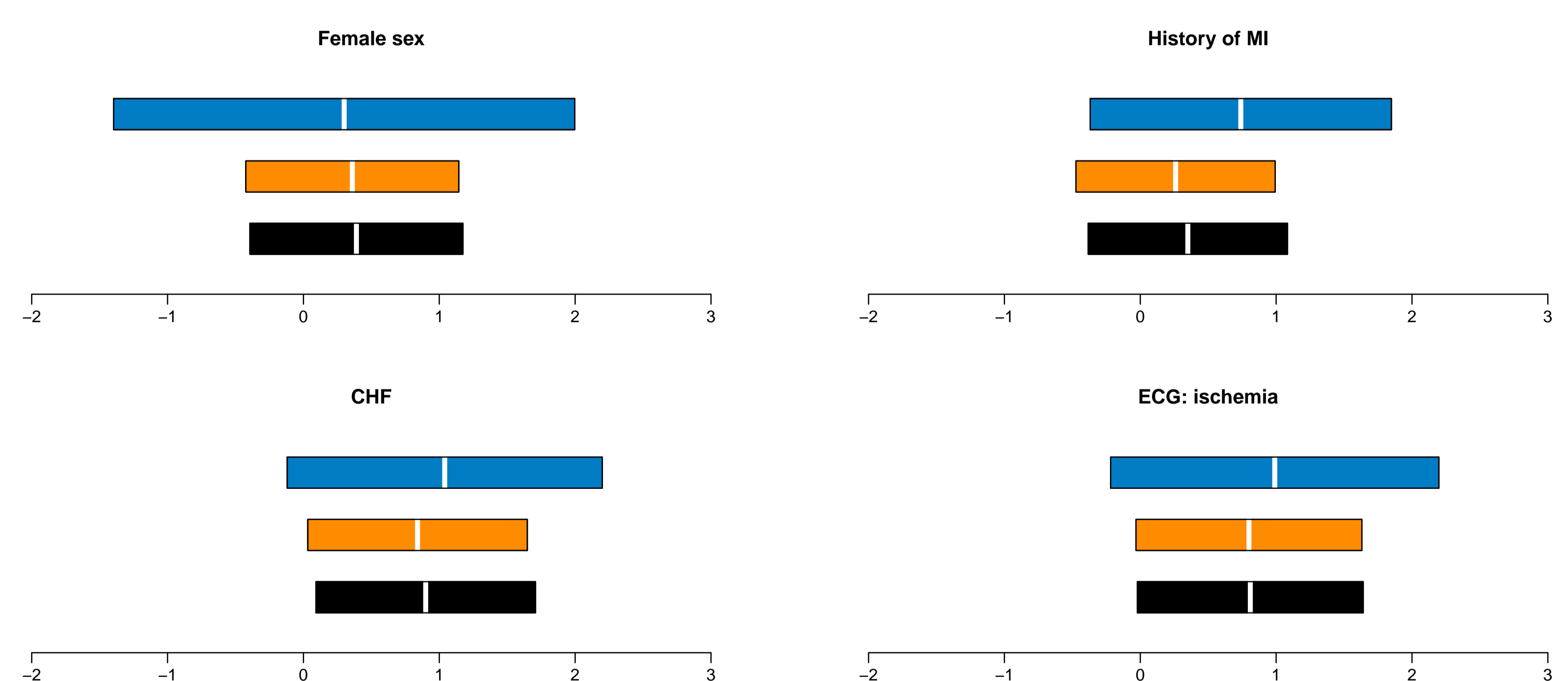
### RESULTS SIMULATION STUDY

Mean squared error and coverage of the 95% confidence interval for estimated regression coefficients  $\beta_1$  (for  $\sigma_h = 0$  and  $\sigma_h = 0.20$ )



### RESULTS APPLICATION

Calculation of adapted associations for Mortality of Elective Aortic Aneurysm Surgery



### REFERENCES

1. Gelman A, Jakulin A, Pittau G.M., Su Y.S. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* 2008; 2(4): 1360-1383.
2. Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987; 9: 1-30.
3. Steyerberg EW, Eijkemans MJ, Van Houwelingen JC, Lee KL, Habbema JD. Prognostic models based on literature and individual patient data in logistic regression analysis. *Stat Med* 2000; 19(2): 141-160.