UMC Utrecht
Julius Center

A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes

Thomas Debray, PhD

Julius Center for Health Sciences and Primary Care,
University Medical Center Utrecht, The Netherlands

Cochrane Netherlands, Utrecht, The Netherlands

thomasdebray.be

I have no actual or potential conflict of interest in relation to this presentation.

Introduction

- Abundance of published prediction models to estimate absolute risk in individual patients
  - Cardiovascular disease ($>$ 350 models) doi
  - Traumatic brain injury ($>$ 100 models) doi
  - Breast cancer ($>$ 50 models) doi
- External validation studies are increasingly common
  - Apply published model(s) in new patients
  - Compare predicted and observed outcomes
  - Quantify discrimination and calibration performance
  - TRIPOD guidelines for conduct and reporting

Introduction

- Estimates of prediction model performance are likely to vary
  - Sampling error
  - Differences in predictor effects
  - Differences in patient spectrum
- Interpretation of validation study results often difficult
  - Reproducibility of model predictions
  - Transportability across different settings and populations
- Synthesis of validation studies is needed
  - To assess the model's likely performance in new settings or populations
  - To better understand under what circumstances developed models perform adequately or require further adjustments

Formal guidance for systematic review and meta-analysis

- Debray TPA, et al. A guide to systematic review and meta-analysis of prediction model performance. BMJ 2017.
- Debray TPA, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. Stat Methods Med Res 2018.
- Implementation in R package metamisc

Motivating example

Framingham Risk Score

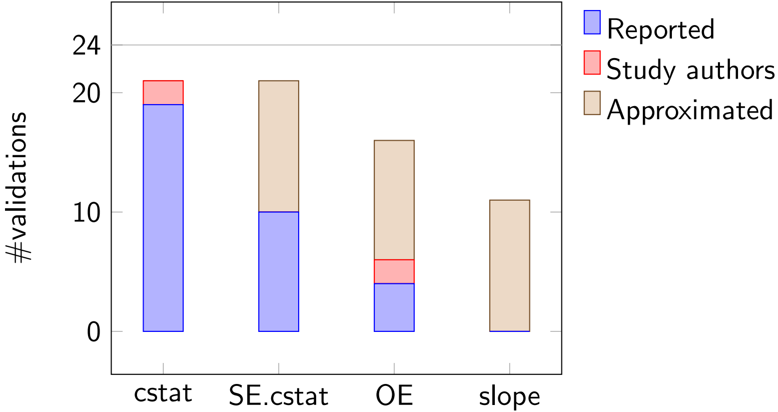
- Model type: Cox regression
- Outcome: Fatal or non-fatal coronary heart disease (CHD)
- Timing: Initial CHD within 10 years
- Evidence: 24 validations in male populations

Summarize estimates of model performance

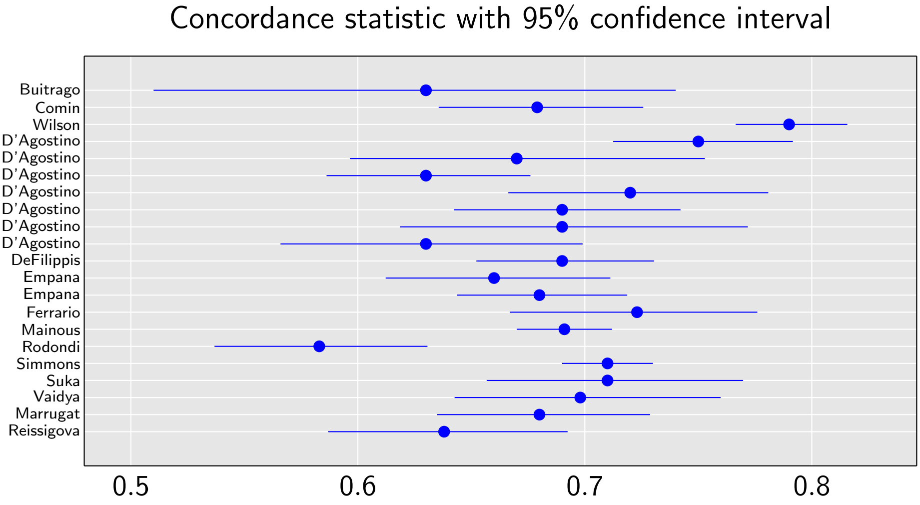
- Concordance statistic (`cstat`)
- Ratio of observed versus expected events (`OE`)
- Calibration slope (`slope`)

Data extraction

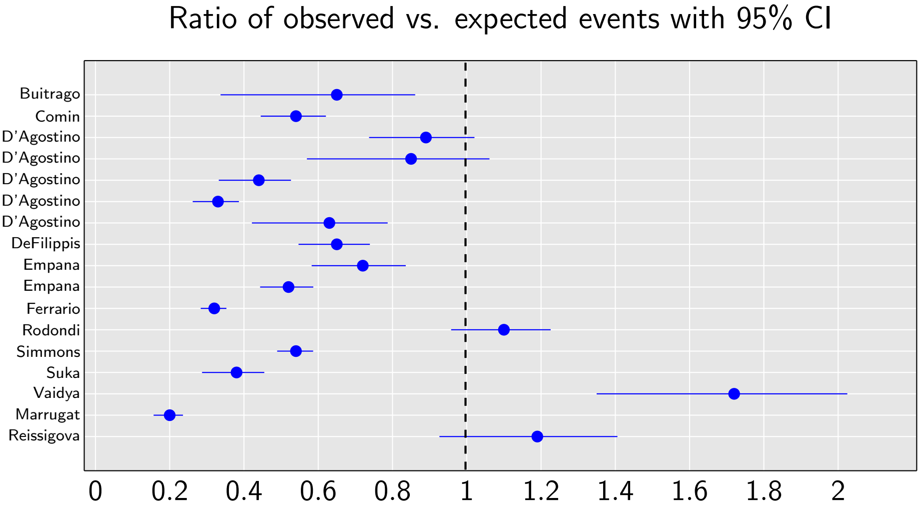
Key problem: Poor and inconsistent reporting of prediction model performance



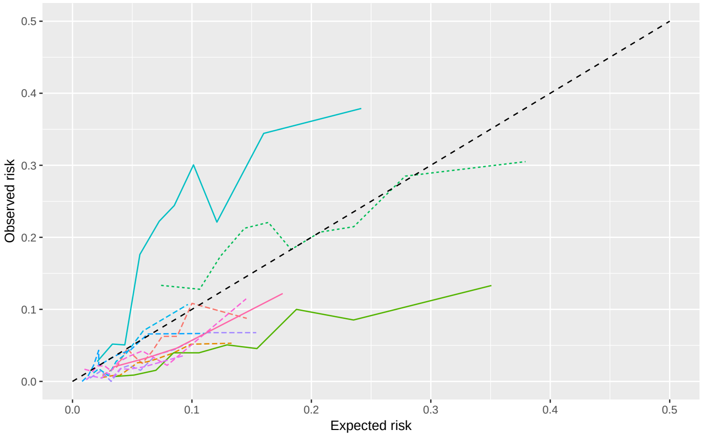
Data extraction



Data extraction



Data extraction



Risk estimates were reported for 5 (dashed lines), 7.5 (dotted lines) and 10 years follow-up (full lines).

Meta-analysis of the c-statistic

- Proposed random effects model

$$\text{logit}(c_i) \sim \mathcal{N}\left(\mu_{\text{discr}}, \text{Var}\left(\text{logit}(c_i)\right) + \tau^2_{\text{discr}}\right)$$

- Weakly informative priors based on empirical data from 26 meta-analyses

Estimation	K	**Summary**	**95% CI**	**95% PI**
REML	21	0.69	0.66 – 0.71	0.59 – 0.77
Bayesian (Unif)	24	0.69	0.66 – 0.71	0.59 – 0.78
Bayesian (Student-t)	24	0.69	0.66 – 0.71	0.59 – 0.78

For 3 studies, we did not have information on c_i but could nevertheless approximate $\text{SE}(c_i)$.

Meta-analysis of the total O:E ratio

- 3 possible random effects models

Option 1 $$\ln(\text{O:E})_i \sim \mathcal{N}\left(\mu_{\text{cal.OE}}, \text{Var}\left(\ln(\text{O:E})_i\right) + \tau^2_{\text{cal.OE}}\right)$$

Option 2 $$O_i \sim \text{Binom}\left(N_i, p_{\text{O},i}\right)$$
$$E_i \sim \text{Binom}\left(N_i, p_{\text{E},i}\right)$$
$$\ln\left(p_{\text{O},i}/p_{\text{E},i}\right) \sim \mathcal{N}\left(\mu_{\text{cal.OE}}, \tau^2_{\text{cal.OE}}\right)$$

Option 3 $$O_i \sim \text{Poisson}\left(E_i \exp(\eta_i)\right)$$
$$\eta_i \sim \mathcal{N}\left(\mu_{\text{cal.OE}}, \tau^2_{\text{cal.OE}}\right)$$

- Weakly informative priors based on empirical data from 16 meta-analyses

Meta-analysis of the total O:E ratio

Estimation	K	Summary	95% CI	95% PI	
REML[1]	6	0.56	0.28 – 1.16	0.09 – 3.62	
Bayesian[1] (Unif)	6	0.61	0.19 – 1.08	0.00 – 2.84	
Bayesian[1] (Student-t)	6	0.61	0.20 – 1.07	0.00 – 2.63	
ML[3]	6	0.56	0.25 – 1.26	0.03 – 11.29	$\star$
Bayesian[3] (Unif)	7	0.60	0.19 – 1.09	0.00 – 2.91	
Bayesian[3] (Student-t)	7	0.60	0.18 – 1.05	0.00 – 2.67	

Meta-analysis of the calibration slope

- Proposed random effects model

$$O_{ij} \sim \mathrm{Binom}(N_{ij}, p_{\mathrm{O},ij})$$
$$\mathrm{logit}(p_{\mathrm{O},ij}) = \alpha_i + \beta_i \mathrm{logit}(P_{\mathrm{E},ij})$$
$$\beta_i \sim \mathcal{N}(\mu_{\mathrm{cal.slope}}, \tau^2_{\mathrm{cal.slope}})$$

Estimation	K	**Summary**	**95% CI**	**95% PI**
ML	3	1.03	0.90 – 1.16	0.20 – 1.87
Bayesian†	3	1.05	0.47 – 1.64	-0.01 – 2.22
Bayesian‡	3	1.05	0.51 – 1.65	-0.06 – 2.17

Final remarks

Meta-analysis of prediction model studies is ...

- Necessary (inferring on generalizability)
- Feasible (methods, guidance and software available)
- Supported (Cochrane Prognosis Methods Group)