Aggregating published prediction models with individual participant data

Thomas Debray

Koffijberg H, Vergouwe Y, Nieboer D, Steyerberg EW, Moons KGM

University Medical Center Utrecht

Julius Center

for Health Sciences and Primary Care

Clinical Prediction Modeling

- Explicit diagnosis and prognosis
- Development of clinical prediction models
- Lack of external validation
- Small datasets & poor generalizability

Meta-analysis

- Allow evidence to accumulate through aggregation
- Extend Model Updating to Model Aggregation
- Break cycle of under-powered derivation, poor generalizability and derivation again.

Challenges

- Lack of data
- Heterogeneity of populations & models
- Average population versus population of interest
- Need for explicit summary models

Bear, as I can, I must, knowing the might of strong Necessity is unconquerable. But touching my fate silence and speech alike are unsupportable.

–Aeschylus, Prometheus Bound

Proposed paradigm

1. Literature review
   - Identify useful models
   - Critical appraisal
2. Validation study
   - Discrimination
   - Calibration
3. Model Updating
   - Adjust literature models to validation sample
   - Update intercept, slope(s)
   - Avoid overfitting (simple updating strategies are preferred)
4. Aggregation of updated models
   - Model Averaging or Stacked Regressions
   - Account for updating complexity

Model Averaging

- Derive probabilistic weights for literature models
$$w_m = \frac{\exp(-0.5\,\mathrm{BIC}_m)}{\sum_{l=1}^{M}\exp(-0.5\,\mathrm{BIC}_l)}, \ \ \mathrm{BIC}_j = -2\,\ell_j + k_j \ln(N_{\mathrm{VAL}})$$

- Average model predictions
$$\overline{p}_i = \sum_{m=1}^{M} w_m \mathcal{M}_m(\mathbf{x})$$

- Estimate summary model
$$\mathrm{logit}(\overline{p}_i) = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ik} + \epsilon_i$$

Properties

- Accounts for model fit and complexity updating method (e.g. $k = 2$ for update intercept and slope)

- Allows implementation of variable selection algorithms

- Explicit summary model

Stacked Regressions

- Weight predictions from the literature models
- Minimize $\sum_i^n (y_i - \alpha_0 - \sum_m^M \alpha_m \mathcal{M}_m(\mathbf{x}))^2$ in the IPD
- Non-negative constraints on the regression slopes α_m
- Update complexity term k_m can be included
- Explicit summary model

# Illustration: predicting Deep Vein Thrombosis

- Validation and updating of 5 literature models ($N = 1,028$):
  - AUC $0.67 \rightarrow 0.70$ ($k = 7$) - Hamilton
  - AUC $0.76 \rightarrow 0.82$ ($k = 10$) - Wells
  - AUC $0.77 \rightarrow 0.83$ ($k = 11$) - Modified Wells
  - AUC $0.81 \rightarrow 0.81$ ($k = 2$) - Gagne
  - AUC $0.82 \rightarrow 0.82$ ($k = 2$) - Oudega
- Model Aggregation
  - Model Averaging $\rightarrow$ **AUC = 0.82**
    Oudega ($w = 0.998$) and Gagne ($w = 0.002$)
  - Stacked Regressions $\rightarrow$ **AUC = 0.86**
    Oudega ($w = 0.41$) and Modified Wells ($w = 0.77$), $\alpha_0 = 0.39$
- Additional validation studies needed!

Simulation Studies

- Available evidence
  - 1 validation sample (14–114 events, 10 predictors, prev 16%)
  - 5 literature models (20 events, backward selection)
- Heterogeneity in literature models
  - Intercept term (A)
  - Intercept term + overall slope (B)
  - Intercept term + overall slope + confounding (C)
- Approaches
  - Model Redevelopment (backward selection, PMLE)
  - Model Updating (intercept + overall slope)
  - Model Aggregation (model averaging, stacked regressions)
- External validation (1,000 events)

Simulation Studies

Re-development with backward selection (line), Re-development with PMLE (dash), Model Updating (dot), Model Averaging and Stacked Regressions.

Simulation Studies

Re-development with backward selection (line), Re-development with PMLE (dash), Model Updating (dot), Model Averaging and Stacked Regressions.

scenario C
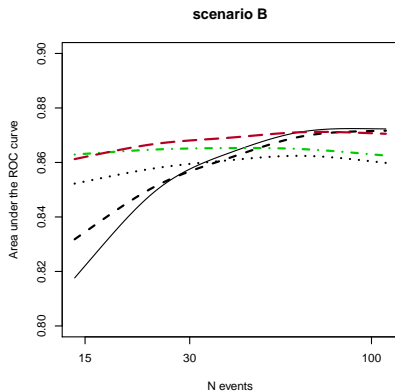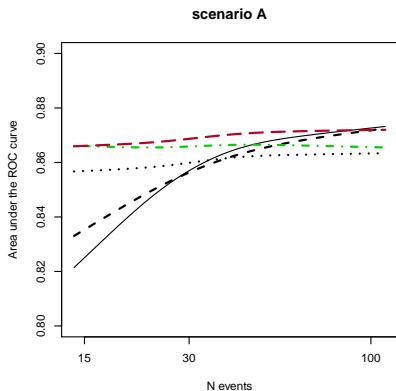
N events
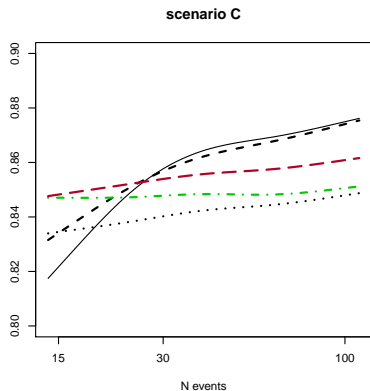
Discussion

Strenghts

- Accumulation of evidence
- No large sample sizes required
- Fairly robust against heterogeneity and overfitting

Extensions

- Variable selection
- Penalized weights (Stacked Regressions)
- Alternative weight schemes (Model Averaging)