Meta-analysis and aggregation of multiple published prediction models

Thomas Debray

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

Supported by Netherlands Organisation for Scientific Research (TOP 9120.8004, 918.10.615 and 916.11.126)

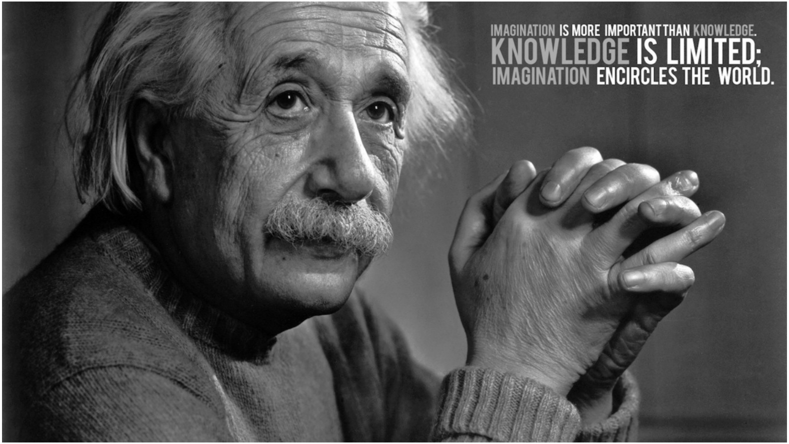
Clinical Prediction Modeling

Model development

- Diagnostic & prognostic outcomes
- Small datasets & overoptimism
- Inappropriate modeling strategies
- Lack of external validation
  $\rightarrow$ model redevelopment
- **Abundance of similar models with poor generalizability**

Evidence aggregation

- Model updating
- IPD meta-analysis
- **Combine prediction models**


IMAGINATION IS MORE IMPORTANT THAN KNOWLEDGE.
KNOWLEDGE IS LIMITED;
IMAGINATION ENCIRCLES THE WORLD.

Evidence aggregation: challenges

- Heterogeneity
  (populations, study designs, model specification, . . .)
- Target population
  (difficult to define without participant data)
- Fully parametric models
  (enhances interpretation & facilitates future implementation)

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

Case study

Diagnosis of **Deep Vein Thrombosis** (DVT)

- Previously published prediction models
  - Wells, Modified Wells (secondary care; rule)
  - Hamilton (secondary care; rule)
  - Gagne, (primary care)
  - Oudega (primary care)
- Validation dataset ($N = 1028$, primary care)

ARTICLE

The Wells Rule Does Not Adequately Rule Out Deep Venous Thrombosis in Primary Care Patients

Ruud Oudega, MD; Arno W. Hoes, MD, PhD; and Karel G.M. Moons, PhD

Background: Using data from secondary care outpatients, Wells and colleagues developed a diagnostic rule to estimate the prob-

testing. Repeated leg ultrasonography was the reference standard to determine the true presence or absence of DVT.

Classical Paradigm

1. Literature search
   Wells, Modified Wells, Hamilton, Gagne, Oudega
2. Critical appraisal
   discard secondary care models?
3. External validation
   identify best models (Oudega & Gagne)
4. Model updating
   intercept update, **logistic calibration**, model revision
5. Recommendations
   use (updated) Oudega model?

No accumulation of other potentially useful models!

Model Averaging (MA)

1. Update literature models
2. Derive probabilistic weights for literature models to average their predictions
   $w_m = \exp(-0.5\ \mathrm{BIC}_m) / \sum_{l=1}^{M} \exp(-0.5\ \mathrm{BIC}_l)$
3. Estimate summary model
   $\mathrm{logit}(\overline{p}_i) = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ik} + \epsilon_i$

Case study: $w_1 = 0.998$ (Oudega), $w_2 = 0.002$ (Gagne)
AUC meta-model = **0.82**

Allows implementation of variable selection algorithms
Explicit summary model

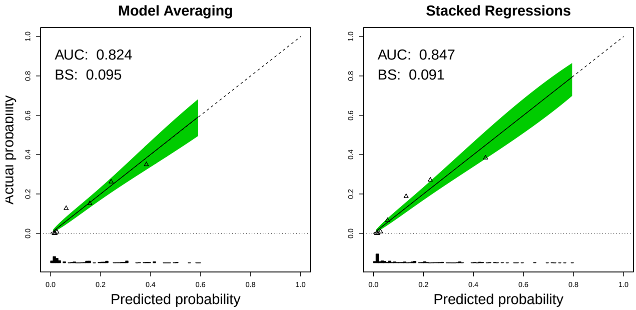
Stacked Regressions (SR)

- Simultaneously updates, discovers and estimates the best combination of literature models
- Minimize $-[\sum_{i=1}^{N} y_i \ln(1 + \exp(-\alpha_0 - \sum_{m=1}^{M} \alpha_m \mathrm{LP}_{im})) - (1 - y_i) \ln(1 + \exp(\alpha_0 + \sum_{m=1}^{M} \alpha_m \mathrm{LP}_{im}))]$
- Non-negative constraints on the regression slopes α_m
- Inspect collinearity! (Variance inflation factor)

Case study: $\alpha_1 = 0.537$ (Oudega), $\alpha_2 = 0.497$ (Gagne) and $\alpha_0 = 1.01$. AUC meta-model: **0.85**

Explicit summary model

Results case study



Meta-model includes 10 predictors (out of 14 possible predictors)
Secondary care models excluded for MA and SR!

Closing remarks

Extension of Model Validation and Updating

- Validity meta-model
- Predictor codings & nonlinearity terms
- Time-to-event data

Advantages

- Parsimonious optimization
- Customizability
- Model weighting (rather than selection)
- Identification of important predictors