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
  → 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
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

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 probability of deep vein thrombosis (DVT). The rule was based on data obtained from previous studies. The study aimed to validate the diagnostic accuracy of the Wells rule in primary care patients, a population for whom the rule was not originally intended.

Methods:
- The study included 1295 consecutive patients who consulted their primary care physician about symptoms suggestive of DVT.
- Patients underwent history-taking and physical examination to calculate the Wells rule score, and D-dimer testing.
- Participants were further work-up by using a combination of symptoms and signs to discriminate among patients with a low, moderate, and high probability of having DVT.
- Measurements: Wells, Modified Wells (secondary care; rule), Hamilton (secondary care; rule), Gagne, (primary care), and Oudega (primary care).

Results:
- Of 1295 consecutive patients who consulted their primary care physician about symptoms suggestive of DVT, 3% were referred for additional, more burdensome, and costly tests in secondary care.
- The Wells rule in primary care patients and compare it with secondary care outpatient clinics. Although it is often argued that secondary care outpatients are similar to primary care patients, differences may exist because of the referral mechanism of primary care physicians. The true prevalence of DVT was never been formally validated in primary care patients, whereas the prevalence was 0.9% in the original study.

Conclusion:
- The Wells rule does not guarantee accurate estimation of risk in primary care patients in whom DVT is suspected. A validation study is needed 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
   \[ w_m = \exp(-0.5 \text{BIC}_m) / \sum_{l=1}^{M} \exp(-0.5 \text{BIC}_l) \]
3. Estimate summary model
   \[ \text{logit}(\bar{p}_i) = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ik} + \epsilon_i \]

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 
  \[-\left[ \sum_{i=1}^{N} y_i \ln(1 + \exp(-\alpha_0 - \sum_{m=1}^{M} \alpha_m LP_{im})) - (1 - y_i) \ln(1 + \exp(\alpha_0 + \sum_{m=1}^{M} \alpha_m LP_{im})) \right]\]
- Non-negative constraints on the regression slopes \(\alpha_m\)
- 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