UMC Utrecht

Meta-analysis of clinical prediction models

Thomas P.A. Debray, PhD

Introduction

"*We demand rigidly defined areas of doubt and uncertainty!*"
Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

UMC Utrecht

Clinical Prediction Models

- Estimation of absolute risk using multiple predictors
  - Demographic characteristics
  - Clinical history and physical examination
  - Medical imaging, elekrofysiology, pathology
  - Biomarker tests
- Diagnostic prediction models
  - Predict presence of a certain disease or condition
- Prognostic prediction models
  - Predict future occurrence of a certain outcome

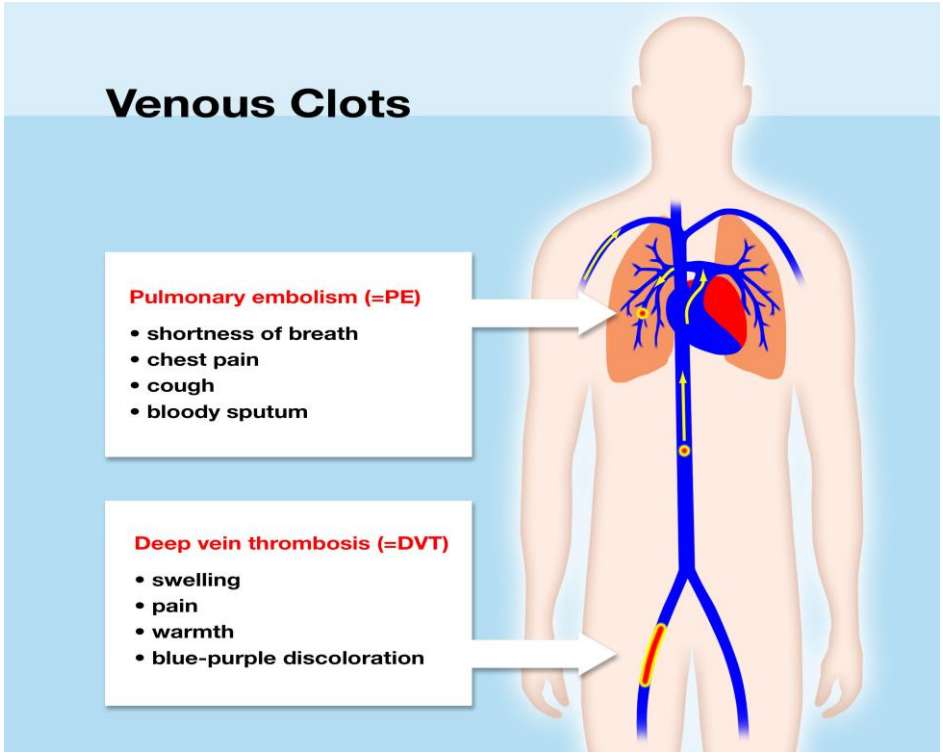
- Guide healthcare professionals and individuals

Clinical Prediction Models

- Developed from subject-level data
  - Cross-sectional studies (diagnostic models)
  - Prospective cohort study (prognostic models)
  - (Other designs)
- Statistical data analyses
  - Data cleaning
  - Predictor selection
  - Missing data
- Presentation
  - Equations to calculate outcome risk
  - Score charts

Deep Venous Thrombosis (DVT)

- Blood clot that forms in a vein in the body (lower leg/thigh)
- If blood clot breaks off -> blood stream -> lungs -> blockage
- Pulmonary embolism, preventing oxygenation of blood
- Potentially causing death



Diagnosis of DVT

- Limited value of signs and symptoms in primary care
- Most patients suspected of DVT referred to secondary care
- Reference standard: ultrasonography (CUS)
- Burden on patients and health care budgets

Need for developing **multivariable prediction models**

- Predict presence of DVT in suspected patients
  - Patient history and physical examination
  - Biomarker test results: D-dimer test
- Primary care versus secondary care

Diagnosis of DVT

WELLS Score (DVT) (*)

• Active cancer (treatment ongoing or within previous 6 months, or palliative treatment)	1
• Paralysis, paresis, or recent plaster immobilization of the lower extremities	1
• Recently bedridden for 3 days or more, or major surgery within the previous 12 weeks requiring general or regional anesthesia	1
• Localized tenderness along the distribution of the deep venous system	1
• Entire leg swollen	1
• Calf swelling > 3 cm compared to asymptomatic leg (measuring 10 cm below tibial tuberosity)	1
• Pitting edema confined to the symptomatic leg	1
• Nonvaricose collateral superficial veins	1
• Previously documented DVT	1
• Alternative diagnosis at least as likely as DVT	-2

≤ 0 : LOW pretest probability
1 or 2 : MODERATE pretest probability
≥ 3 : HIGH pretest probability

Wells PS, et al. N Engl J Med 2003; 349: 1227-35
Anderson DR, et al. J Thromb Haemost 2003; 1: 645-51

(*) In patients with symptoms in both legs, the more symptomatic leg is used

Diagnosis of DVT

Characteristics	Hamilton	Modified Wells
Plaster immobilization of lower limb	2	1
Active malignancy (within 6 months or current)	2	1
Strong clinical suspicion of deep venous thrombosis by the emergency physicians without other diagnostic possibilities	2	–
Bed rest (>3 days) or recent surgery (within 4 weeks)	1	1
Male sex	1	–
Calf circumference >3 cm on affected side (measured 10 cm below tibial tuberosity)	1	1
Erythema	1	–
Localized tenderness along the distribution of the deep venous system	–	1
Entire leg swollen	–	1
Pitting edema confined to the symptomatic leg	–	1
Collateral superficial veins (nonvaricose)	–	1
Previously documented deep vein thrombosis	–	1
Alternative diagnosis at least as likely as deep vein thrombosis	–	−2
Unlikely versus likely cutoff score	2 or less	1 or less

Diagnosis of DVT

Diagnostic variables	Odds ratio	Regression coefficient*	p-value	Points for the rule
Male gender	1.80 (1.36 – 2.16)	0.59	<0.001	1
Oral contraceptive use	2.12 (1.32 – 3.35)	0.75	0.002	1
Presence of malignancy	1.52 (1.05 – 2.44)	0.42	0.082	1
Recent surgery	1.46 (1.02 – 2.09)	0.38	0.044	1
Absence of leg trauma	1.82 (1.25 – 2.66)	0.60	0.002	1
Vein distension	1.62 (1.19 – 2.20)	0.48	0.002	1
Calf difference ≥ 3 cm	3.10 (2.36 – 4.06)	1.13	<0.001	2
D-dimer abnormal	20.3 (8.25 – 49.9)	3.01	<0.001	6
Constant		–5.47		

DVT= deep vein thrombosis; *=natural logarithm of the odds ratio; D-dimer abnormal for VIDAS ≥ 500 ng/ml and Tinaquant ≥ 400 ng/ml. Probability of DVT as estimated by the final model =1/(1+exp-(-5·47 + 0·59*male gender + 0·75*OC use + 0·42*presence of malignancy + 0·38*recent surgery + 0·60*absence of leg trauma + 0·48*vein distension + 1·13*calf difference ≥ 3cm + 3·01*abnormal D-dimer)).

Diagnosis of DVT

TABLEAU II

Analyse multivariée : modèle de régression logistique final prédisant la présence d'une thrombose veineuse profonde

Variable	p	odds ratio	coefficient
Immobilisation médicale dans le mois précédent (alitement > 48 h ou paralysie)	0,07	1,9 (1,0–3,7)	0,62
Contraception oestroprogestative	0,02	4,0 (1,2–12,9)	1,38
Antécédent personnel de MVTE	0,02	2,1 (1,1–4,0)	0,74
Cancer évolutif	<0,01	7,3 (2,4–22,1)	1,99
Diminution du ballant du mollet	0,01	2,3 (1,3–4,1)	0,83
Diagnostic alternatif au moins aussi probable	<0,01	0,1 (0,1–0,3)	−2,08

External validation

- Which model should we use?
- What performance can we expect?
- Does the model require improvements/changes?
- Or, should we rather develop a model from scratch?

External validation is needed!

- Identify and evaluate existing models
- Assess performance in a new sample
- Compare predicted probabilities to observed outcomes
- Distinguish between discrimination and calibration

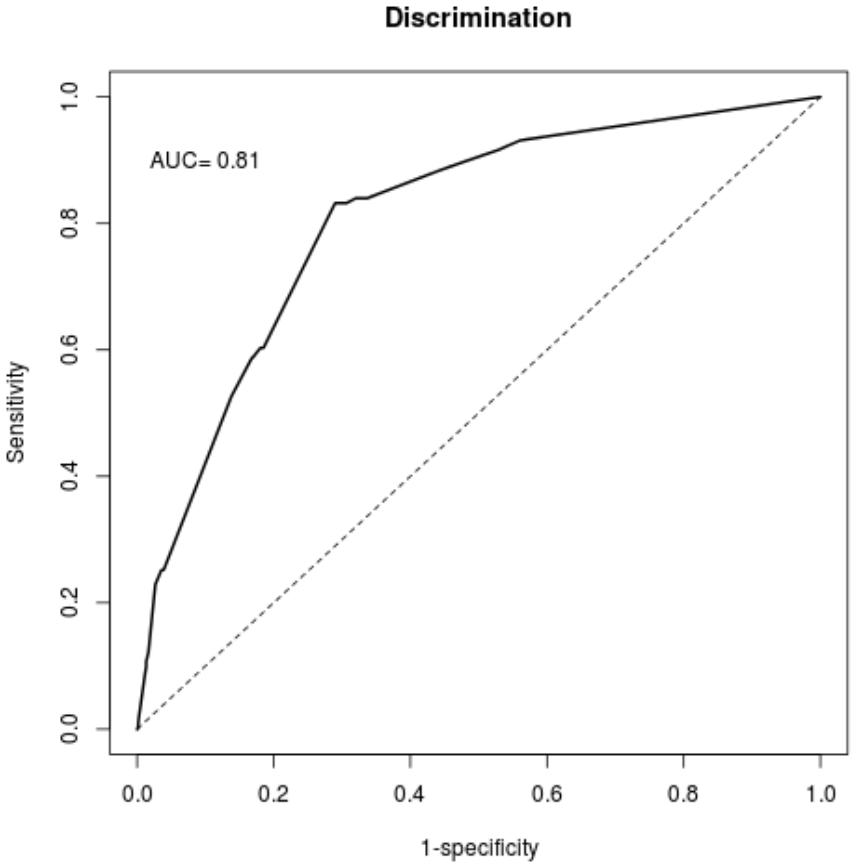
External validation

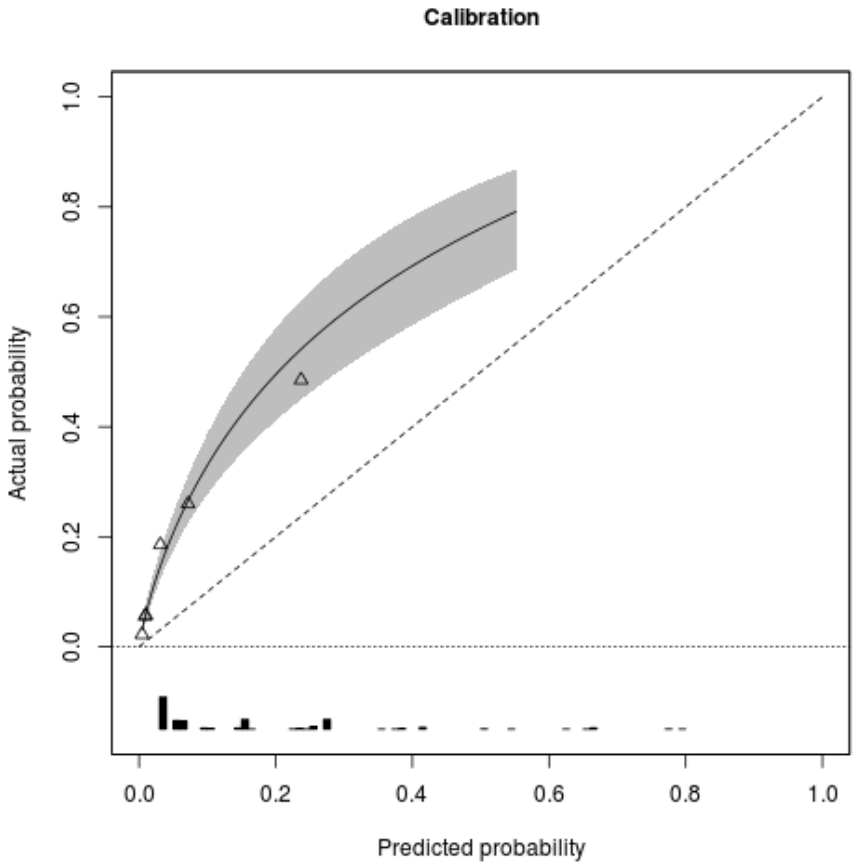
Validation sample for DVT models

- Prospective management study
- 300 primary care practices in 3 regions of the Netherlands (Amsterdam, Maastricht, and Utrecht)
- Outcome: incidence of symptomatic venous thromboembolism during 3-month follow-up
- 1028 patients with clinically suspected DVT
- 131/1028 patients eventually diagnosed with DVT

Question: Can the previously identified models predict which subjects have DVT?

External validation (Gagne)





External validation

- Discrimination secondary care models
  - 0.66 (Hamilton)
  - 0.76 (Wells)
  - 0.77 (modified Wells)
- Discrimination primary care models
  - 0.81 (Gagne)
  - 0.82 (Oudega)

Remark: Secondary-care models may not adequately rule out DVT in primary-care settings!

External validation

Annals of Internal Medicine
Established in 1927 by the American College of Physicians

19 July 2005, Vol 143, No. 2>

< Previous in this Issue Next in this Issue >

Articles | 19 July 2005

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

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

[+] Article and Author Information

Companion Article(s):

Meta-Analysis: The Value of Clinical Assessment in the Diagnosis of Deep Venous Thrombosis

Use of a Clinical Prediction Score in Patients with Suspected Deep Venous Thrombosis: Two Steps Forward, One Step Back?

Ann Intern Med. 2005;143(2):100-107. doi:10.7326/0003-4819-143-2-200507190-00008

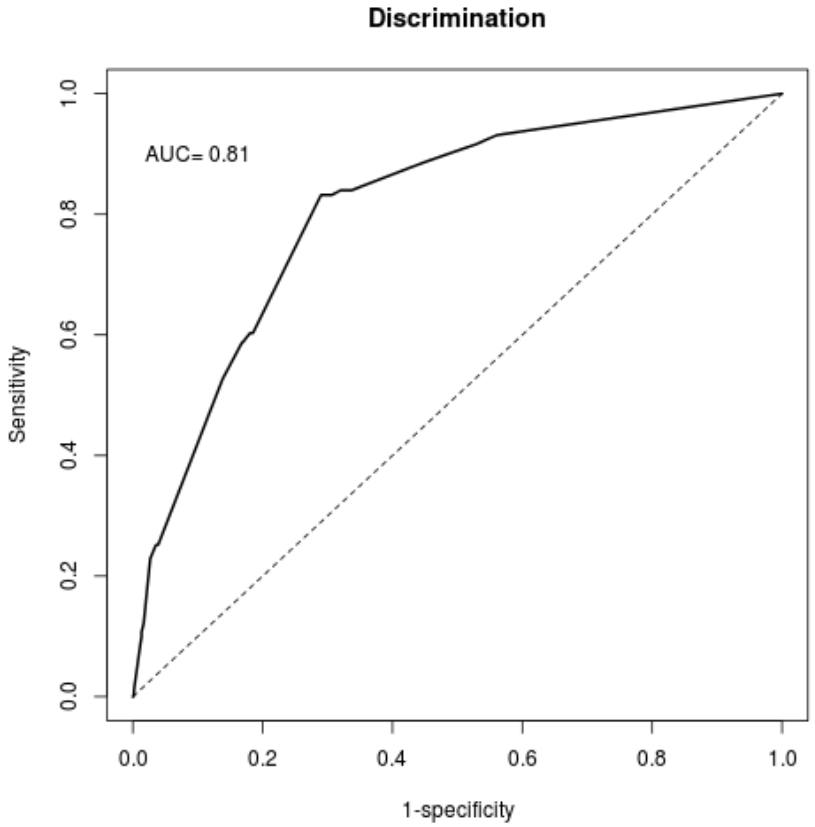
Model updating

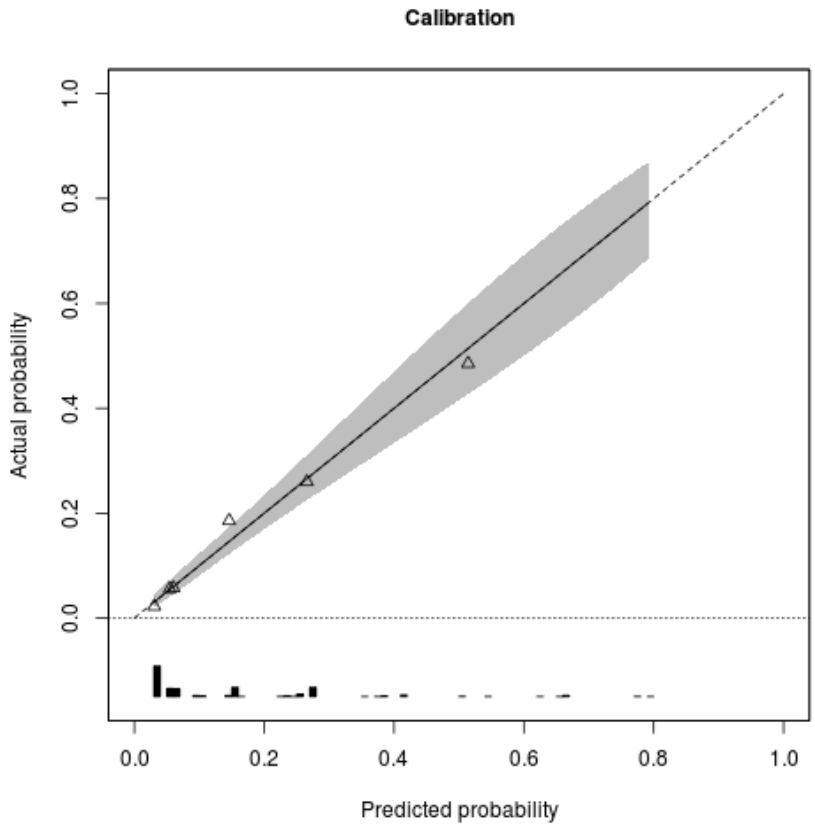
Adjust promising models to the validation sample

- Adjust intercept
  correct for different outcome prevalence
- Adjust intercept and common slope
  correct for different outcome prevalence and predictor effects that are over-optimistic
- More advanced updating procedures
  - Adjust a particular regression coefficient
  - Re-estimate all regression coefficients
  - Add completely new predictors

Remark: updating procedures reduce insight into model validity as new parameters are being estimated

Model updating (Gagne)

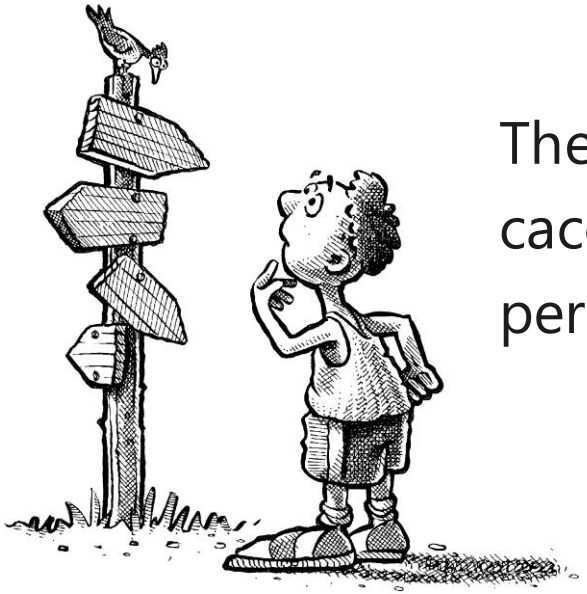




Update of intercept and common slope

Caveats of prediction modeling research

- Most models are never validated
- Model redevelopment versus model updating
- Risk of overfitting
- Prior knowledge not optimally used
- Incompatibility and confusion



The user must typically choose between a cacophony of existing models for which performance may be obsurce

Meta-analysis of prediction models

- Meta-analysis (therapeutic research)
  - Synthesize evidence from multiple trials
  - Obtain a summary estimate of treatment effect
  - Facilitate detailed analyses of effect modification
- Meta-analysis (prognostic research)
  - Synthesize evidence on prognostic factors
  - Aggregate literature models into a meta-model that is optimized for validation sample
  - Improve generalizability of meta-model across different patient populations
- How to combine models with similar predictors?
- How to combine models with different predictors?

Meta-analysis of prediction models

Statistics
in Medicine

Research Article

Aggregating published prediction models with individual participant data: a comparison of different approaches

Thomas P.A. Debray[1,*], Hendrik Koffijberg[1], Yvonne Vergouwe[2], Karel G.M. Moons[1,†] and Ewout W. Steyerberg[2,†]

Article first published online: 26 JUN 2012

DOI: 10.1002/sim.5412

CrossMark click for updates

Issue

Statistics in Medicine
Volume 31, Issue 23, pages 2697–2712, 15 October 2012

SEARCH

In this issue

Advanced > Saved Searches >

ARTICLE TOOLS

Get PDF (179K)
Save to My Profile
E-mail Link to this Article
Export Citation for this Article
Get Citation Alerts
Request Permissions

Share

Additional Information (Show All)

How to Cite | Author Information | Publication History | Funding Information

Meta-analysis of prediction models

Aggregation of prediction models with **similar predictors**

- Identify common predictors
  - restore missing coefficients and standard errors where necessary (imputation)
- Pooling of predictor effects
  - calculate weighted average of regression coefficients
  - account for differences in precision
  - account for heterogeneity across studies
- Meta-model for average or specific study population
  - Relevance of literature versus validation sample
  - Adjust intercept term to local circumstances

Meta-analysis of prediction models

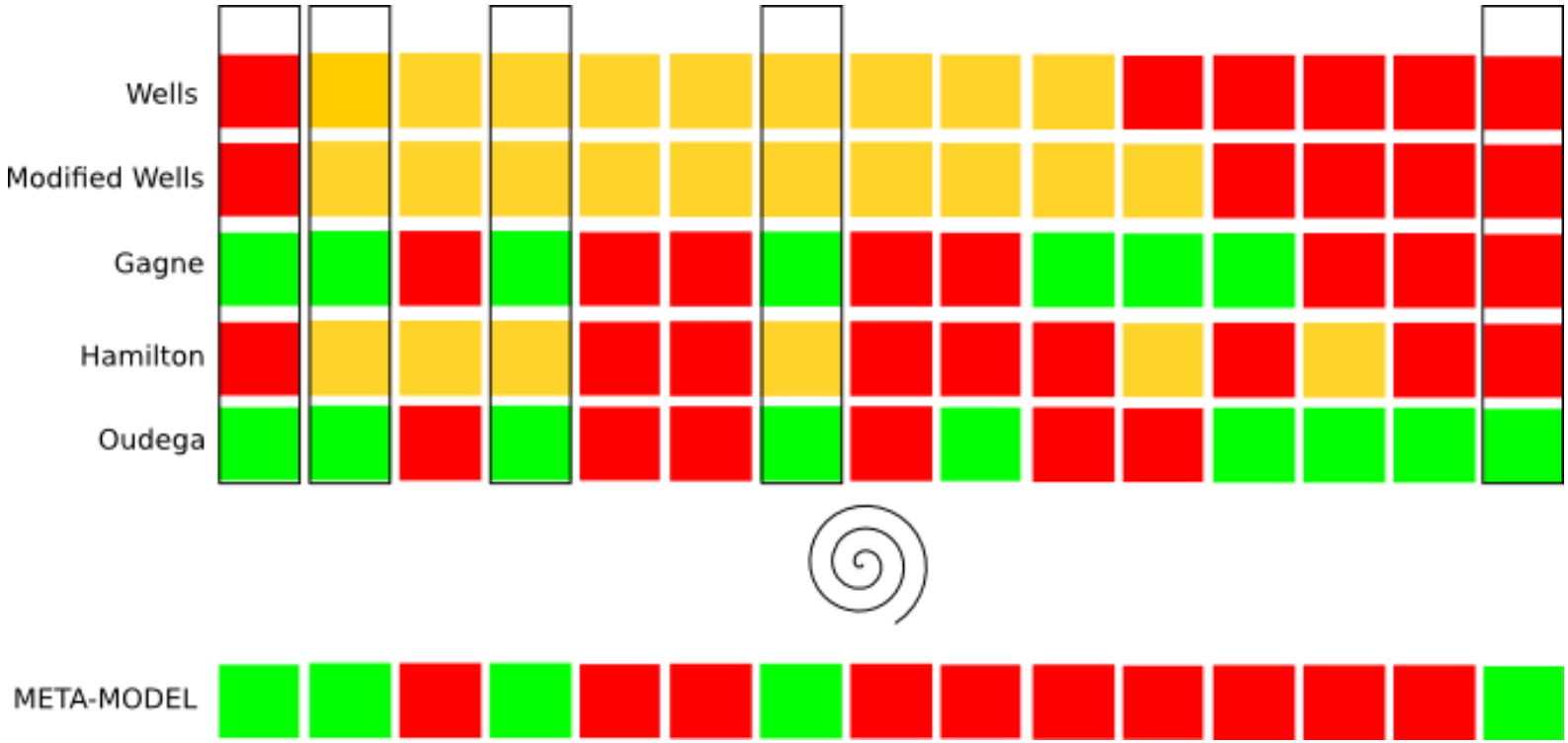
Aggregation of prediction models with **similar predictors**

- Univariate meta-analysis
  - pool predictor effects separately
- Multivariate meta-analysis
  - simultaneous pooling of all predictor effects
- Multivariate meta-analysis + Bayesian inference
  - pooled predictor effects from the literature are used as prior information for the predictor effects in the validation sample

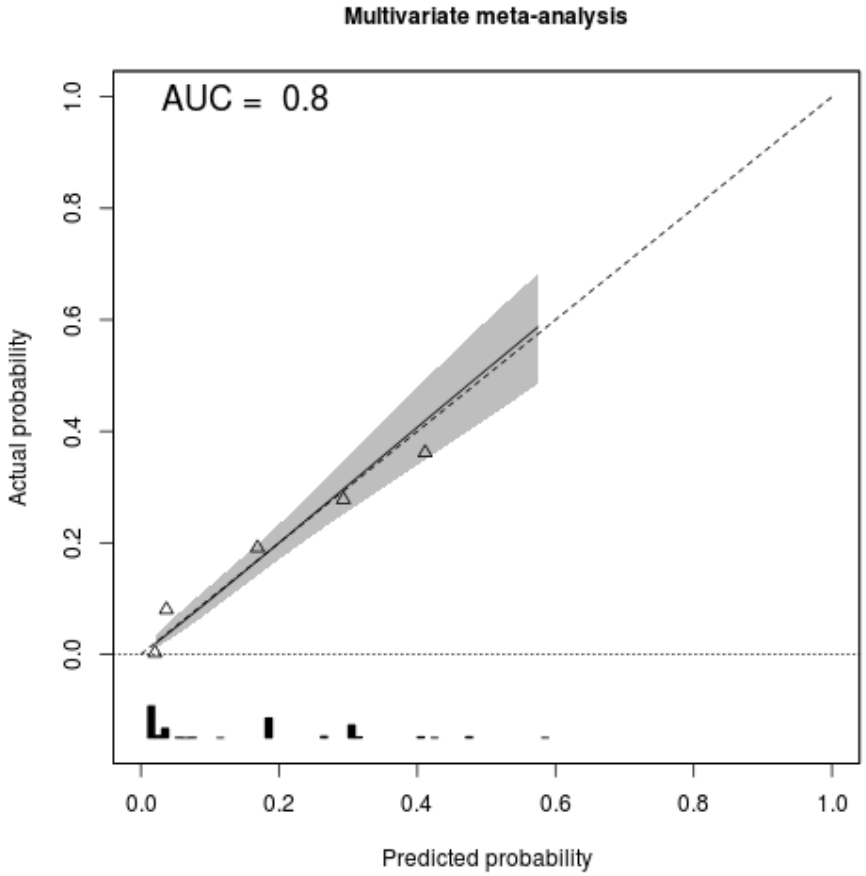
Meta-analysis of prediction models

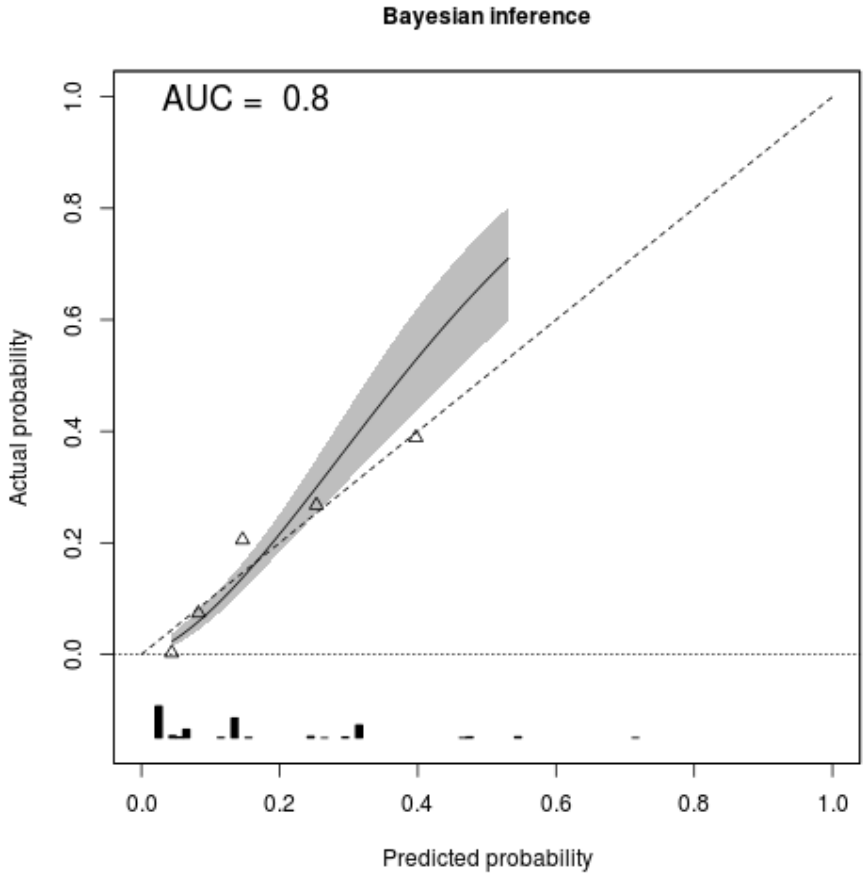
Aggregation of prediction models with **similar predictors**

- Diagnosis of DVT: focus on 4 common core predictors (+ intercept term)



Meta-analysis of prediction models





Meta-analysis of prediction models

Aggregation of prediction models with **similar predictors**

- (Simplified) meta-model
  - fewer predictors
  - adjusted for validation sample (baseline risk)
  - similar performance as best literature model

Implementation difficult when literature models differ much in terms of included predictors

Meta-analysis of prediction models

Statistics
in Medicine

Research Article

Meta-analysis and aggregation of multiple published prediction models

Thomas P.A. Debray[1,*], Hendrik Koffijberg[1], Daan Nieboer[2], Yvonne Vergouwe[2], Ewout W. Steyerberg[2] and Karel G.M. Moons[1]

Article first published online: 14 JAN 2014

DOI: 10.1002/sim.6080

CrossMark click for updates

Issue

Statistics in Medicine

Early View (Online Version of Record published before inclusion in an issue)

Additional Information (Show All)

How to Cite | Author Information | Publication History

Meta-analysis of prediction models

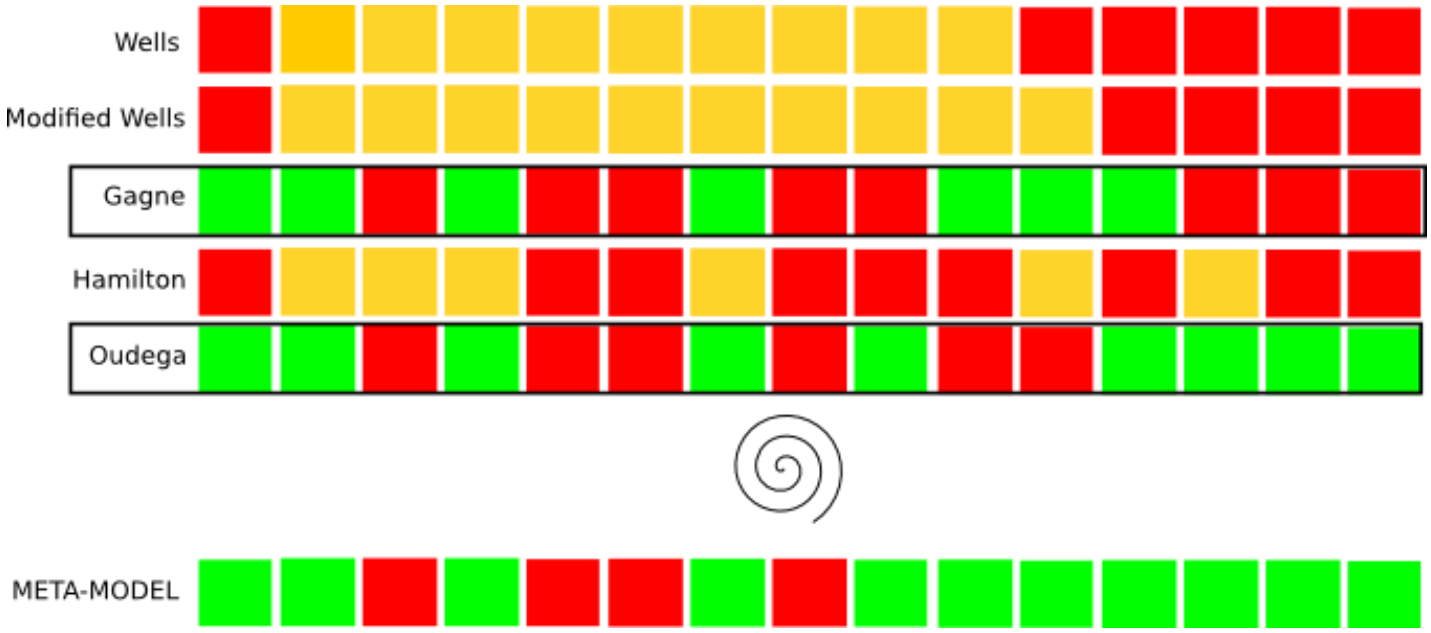
Aggregation of models with **different predictors**

- Model averaging
  1. Update literature models
  2. Calculate predictions for each subject, for each model
  3. Evaluate performance literature models
  4. Calculate weights based on model fit and updating complexity (BIC)
  5. Obtain (weighted) average predictions
  6. Calculate summary model

Meta-analysis of prediction models

Aggregation of models with **different predictors**

- DVT case study
  - update intercept and common slope of all models
  - Weights: **0.998** (Oudega)**, 0.002** (Gagne), 0 (other models)



Meta-analysis of prediction models

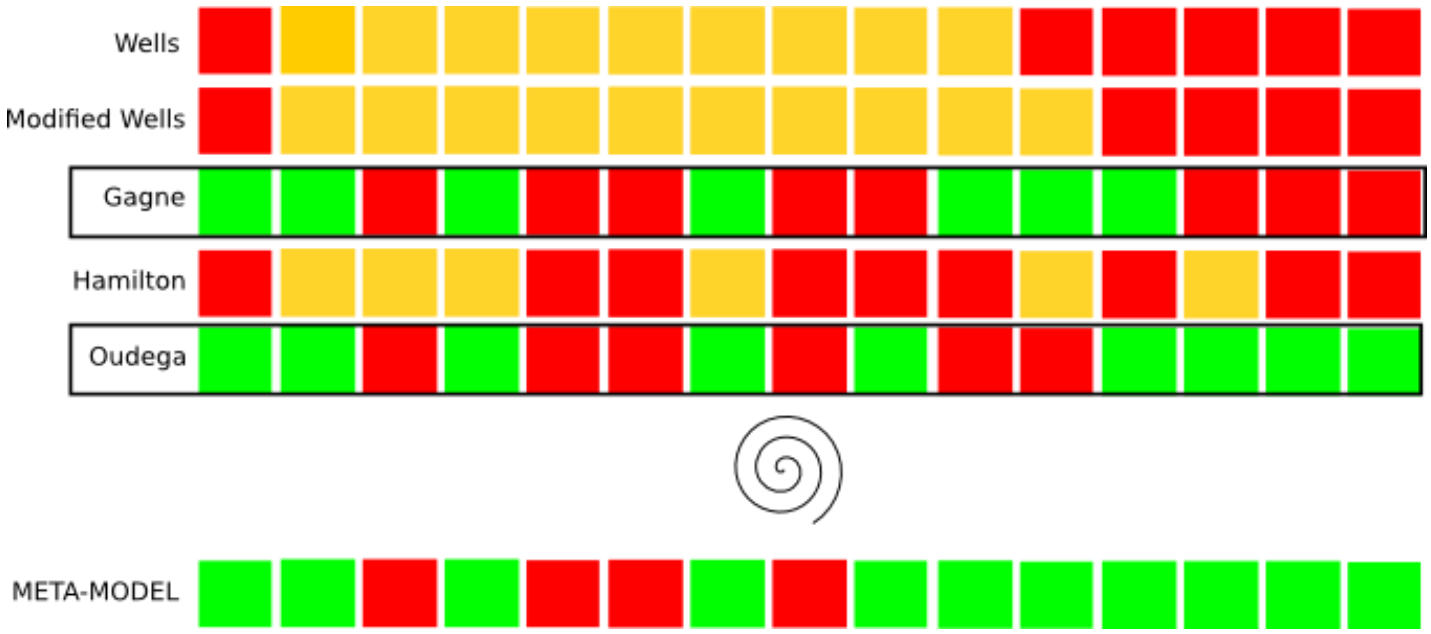
Aggregation of models with **different predictors**

- Stacked regressions
  - Weight predictions from literature models
  - Discard models with little (added) value
  - Update common intercept and overall slope
  - No distinct steps, one straightforward estimation procedure
  - Borrows less information from validation sample (as compared to model averaging)

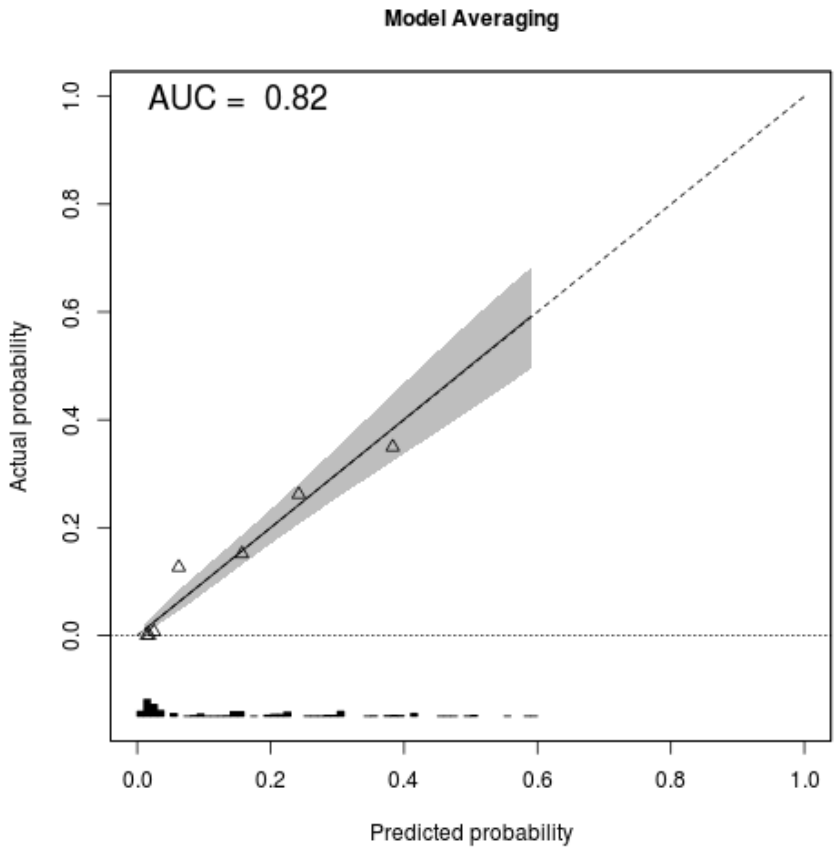
Meta-analysis of prediction models

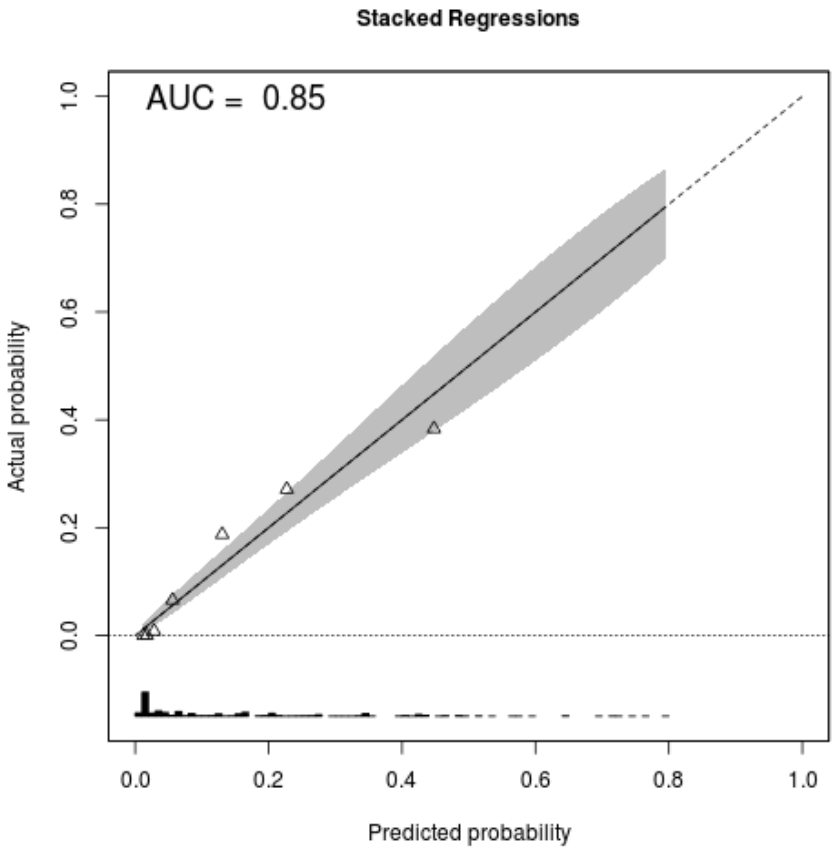
Aggregation of models with **different predictors**

- DVT case study
  - Weights: **1.01** (α), **0.537** (Oudega)**, 0.497** (Gagne), 0 (other models)



Meta-analysis of prediction models





Meta-analysis of prediction models

External validation of meta-models

- Primary Care (N=791)
  - Best literature model: AUC = 0.77, slope = 1.13
  - Model Averaging: AUC = 0.77, slope = 1.13
  - Stacked Regressions: AUC = 0.74, slope = 0.82
- Secondary Care (N=1756)
  - Best literature model: AUC = 0.84, slope =1.29
  - Model Averaging: AUC = 0.86, slope = 1.29
  - Stacked Regressions: AUC = 0.88, slope = 1.33

Meta-model outperforms existing models for primary and secondary care settings!

Meta-analysis of prediction models

Simulation studies

- Model re-development only useful when
  - Large (validation) sample available
  - Literature models too heterogeneous with target population (i.e. differences beyond intercept and common slope)
- For small (validation) samples:
  - Model redevelopment techniques (e.g. backward selection or PMLE) outperformed by meta-analysis
  - Model updating techniques outperformed by meta-analysis

Discussion

- Novel paradigm for model development & validation
- Model aggregation versus selective updating
- Better use of prior knowledge, but only if relevant for target population

- Future research
  - Quality appraisal of literature models
  - Alternative weighting schemes
  - Mixed sources of literature evidence
  - Variable selection