UMC Utrecht
Julius Center

Systematic review & meta-analysis of prediction model studies

Thomas Debray, PhD
Julius Center for Health Sciences and Primary Care
Cochrane Netherlands

Prediction

Estimate the absolute risk in individual patients of …

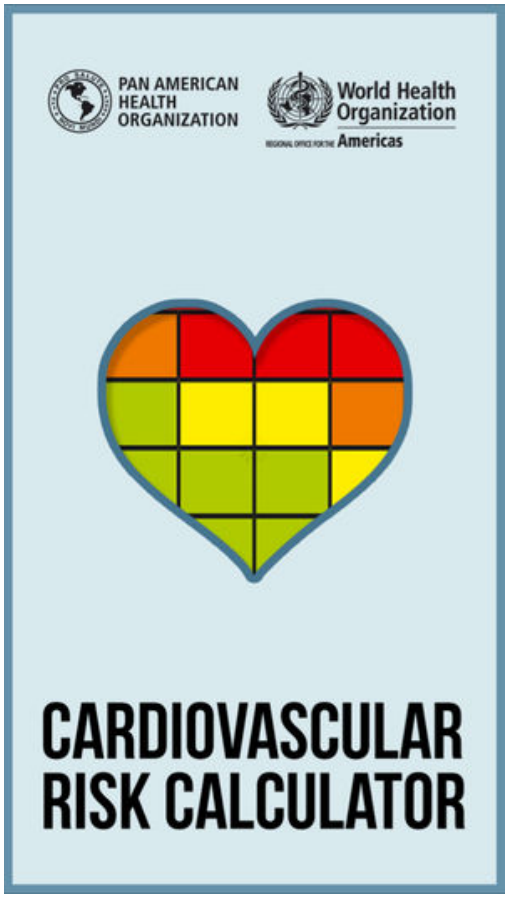
- an outcome's presence (diagnosis)
- an outcome's future occurrence (prognosis)

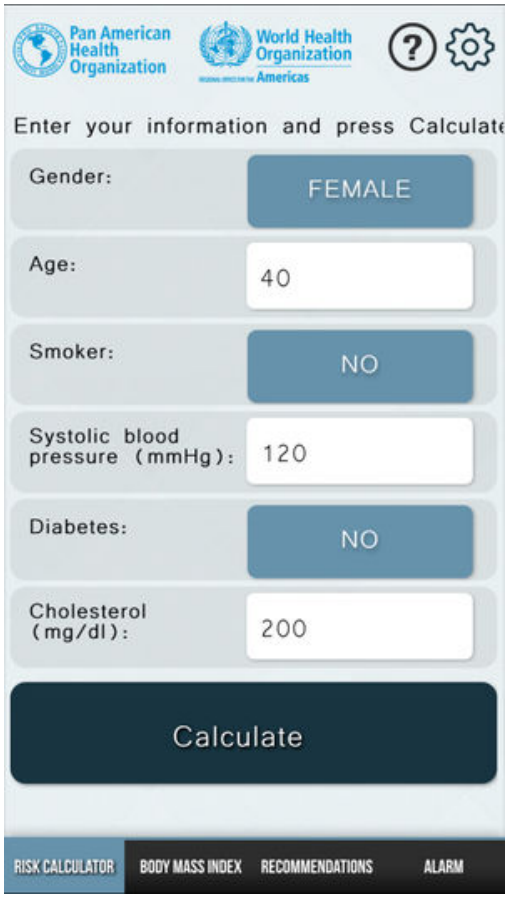
Example

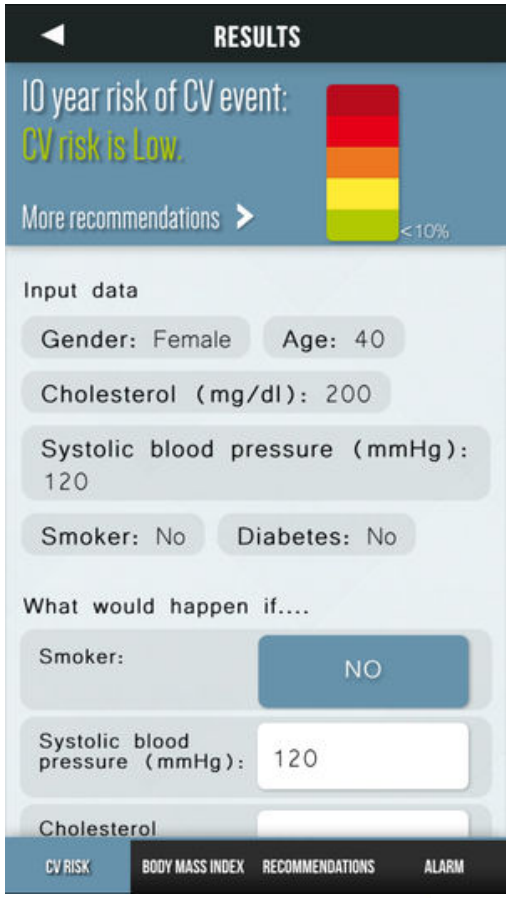
"What is the 10-year risk of cardiovascular disease in a visiting primary care patient?"

Prediction models

Combine information from multiple predictors







Prediction models are abundant

- > 350 models for cardiovascular disease
- > 100 models for brain trauma patients
- > 100 diabetes type 2 models
- > 100 models for prostate cancer
- > 60 models for breast cancer prognosis

The reality

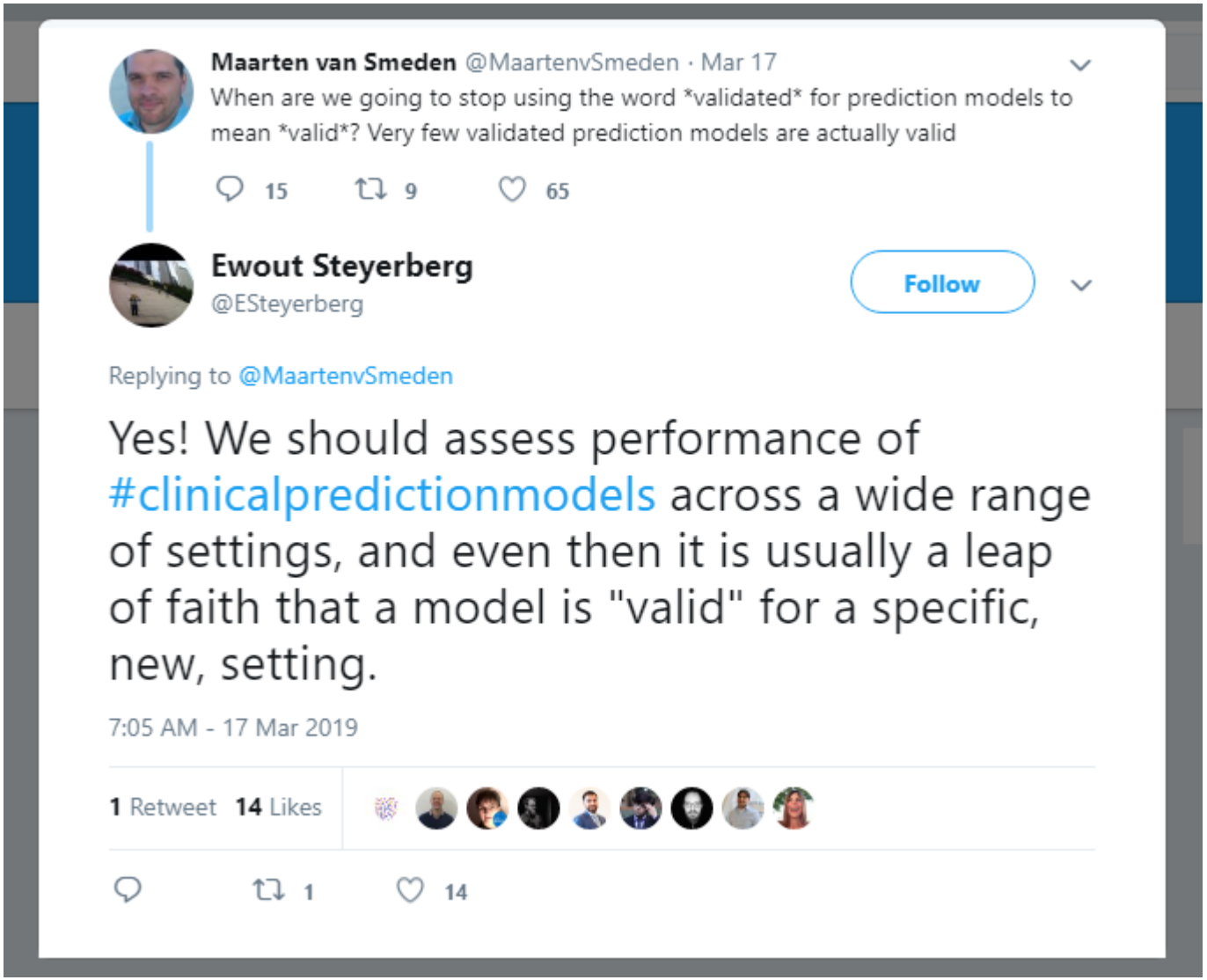
Poor understanding of

- The validity of model predictions in new patients
- The generalizability of prediction models across different settings and populations
- The comparative performance of prediction models
- The clinical impact of prediction models

“All models are wrong, but some are useful”

George Box

The need for evidence synthesis



Maarten van Smeden @MaartenvSmeden · Mar 17

When are we going to stop using the word *validated* for prediction models to mean *valid*? Very few validated prediction models are actually valid

15 9 65

Ewout Steyerberg
@ESteyerberg

Follow

Replying to @MaartenvSmeden

Yes! We should assess performance of #clinicalpredictionmodels across a wide range of settings, and even then it is usually a leap of faith that a model is "valid" for a specific, new, setting.

7:05 AM - 17 Mar 2019

1 Retweet **14** Likes

1 14

The need for evidence synthesis

Synthesis of published prognosis studies may help

- To identify promising markers
  - By summarizing their (incremental) prognostic value
  - By exploring sources of between-study heterogeneity
- To identify promising prediction models
  - By summarizing their predictive performance
  - By exploring generalizability across different settings and populations
  - By evaluating the need for further improvements
- To improve estimation of prediction models
  - By avoiding overfitting in small samples

Summarizing prognosis evidence

Research Methods & Reporting

A guide to systematic review and meta-analysis of prognostic factor studies

BMJ 2019 ; 364 doi: https://doi.org/10.1136/bmj.k4597 (Published 30 January 2019)
Cite this as: *BMJ* 2019;364:k4597

Research Methods & Reporting

A guide to systematic review and meta-analysis of prediction model performance

BMJ 2017 ; 356 doi: https://doi.org/10.1136/bmj.i6460 (Published 05 January 2017)
Cite this as: *BMJ* 2017;356:i6460

Summarizing prognosis evidence

Formal review steps and tools

- Defining the review question (PICOTS)
- Defining the search strategy
- Quantitative data extraction (*)
- Quality appraisal (PROBAST)
- Meta-analysis (*)
- Investigating heterogeneity
- Interpretation (GRADE)
- Reporting (guidelines: REMARK, PRISMA, TRIPOD)

(*) Debray TP et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. Stat Methods Med Res. https://doi.org/10.1177/0962280218785504

Summarizing prognosis evidence

An illustrative example

Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: a systematic review and meta-analysis

Johanna A. Damen, Romin Pajouheshnia, Pauline Heus, Karel G. M. Moons, Johannes B. Reitsma, Rob J. P. M. Scholten, Lotty Hooft & Thomas P. A. Debray

BMC Medicine 17, Article number: 109 (2019) | Cite this article

An illustrative example

PICOTS

- **P**opulation = a general (unselected) population setting
- **I**ntervention = Framingham Wilson 1998
- **C**omparator = Framingham ATP III 2002
- **O**utcome = fatal or nonfatal coronary heart disease
- **T**iming = 10 year
- **S**etting = disease prevention in general population

An illustrative example

Search & identification of eligible studies

- Two previously published systematic reviews
- Search in MEDLINE and Embase
- Citation search in Scopus and Web of Science

Search results: 304 eligible papers

Eligible unique validations with information of the original model's predictive performance:

- Total OE ratio (N = 74)
- Concordance statistic (N = 77)

An illustrative example

Data extraction

- Study design, participant enrolment, study dates
- Population characteristics
- Sample size
- Predictors
- Predicted horizon, predicted outcomes
- Model updating methods
- Model performance (before and after updating)

If relevant information was missing, we contacted the authors and, if unsuccessful, used previously proposed approximations (implemented in R package *metamisc*)

An illustrative example

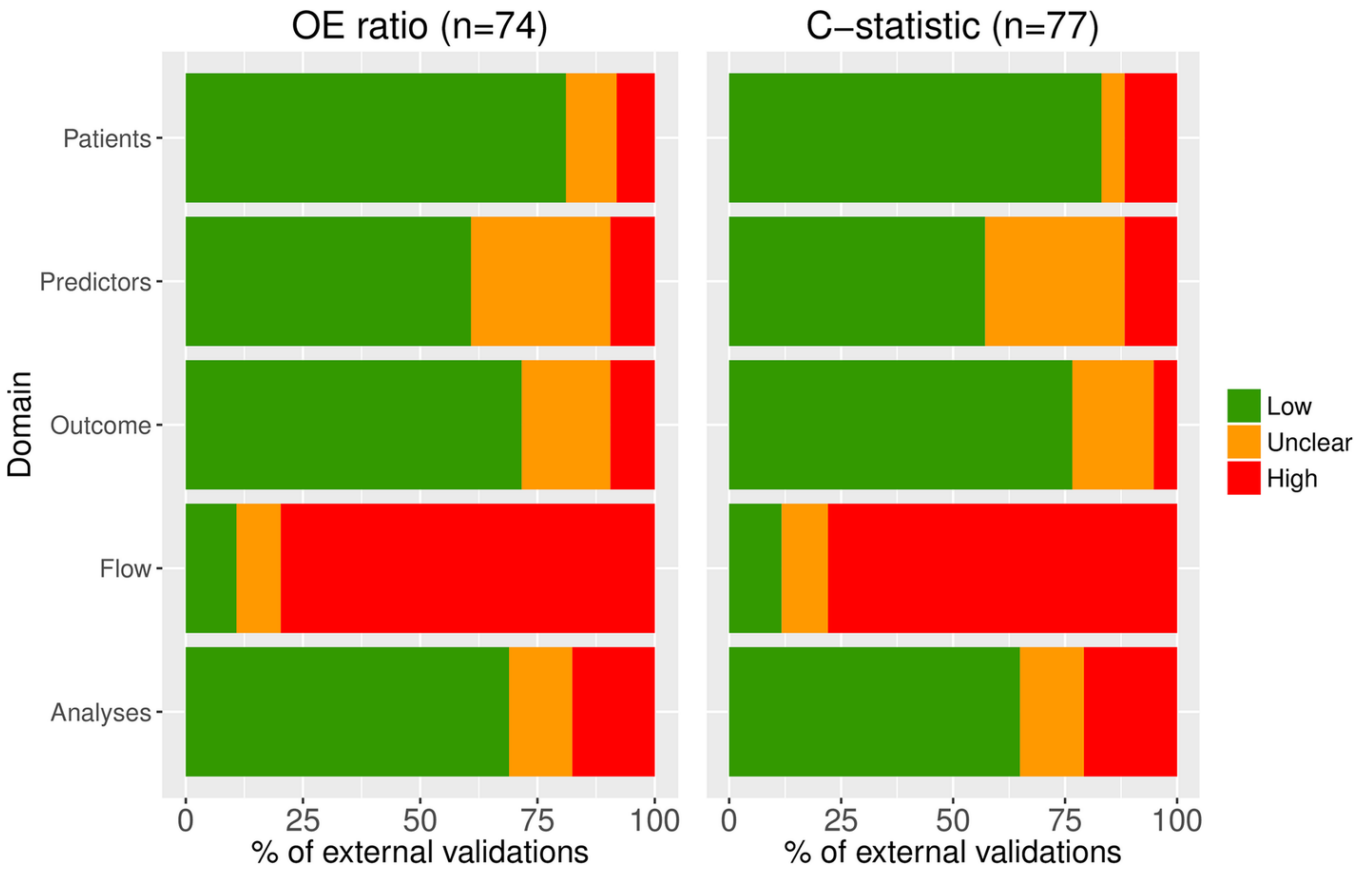
Critical appraisal (PROBAST)

Key findings

- Most validations scored low risk of bias
- Risk of bias for predictors was often unclear due to poor reporting of predictor definitions and measurement methods
- Risk of bias for sample size and participant flow often high due to inadequate handling of missing data

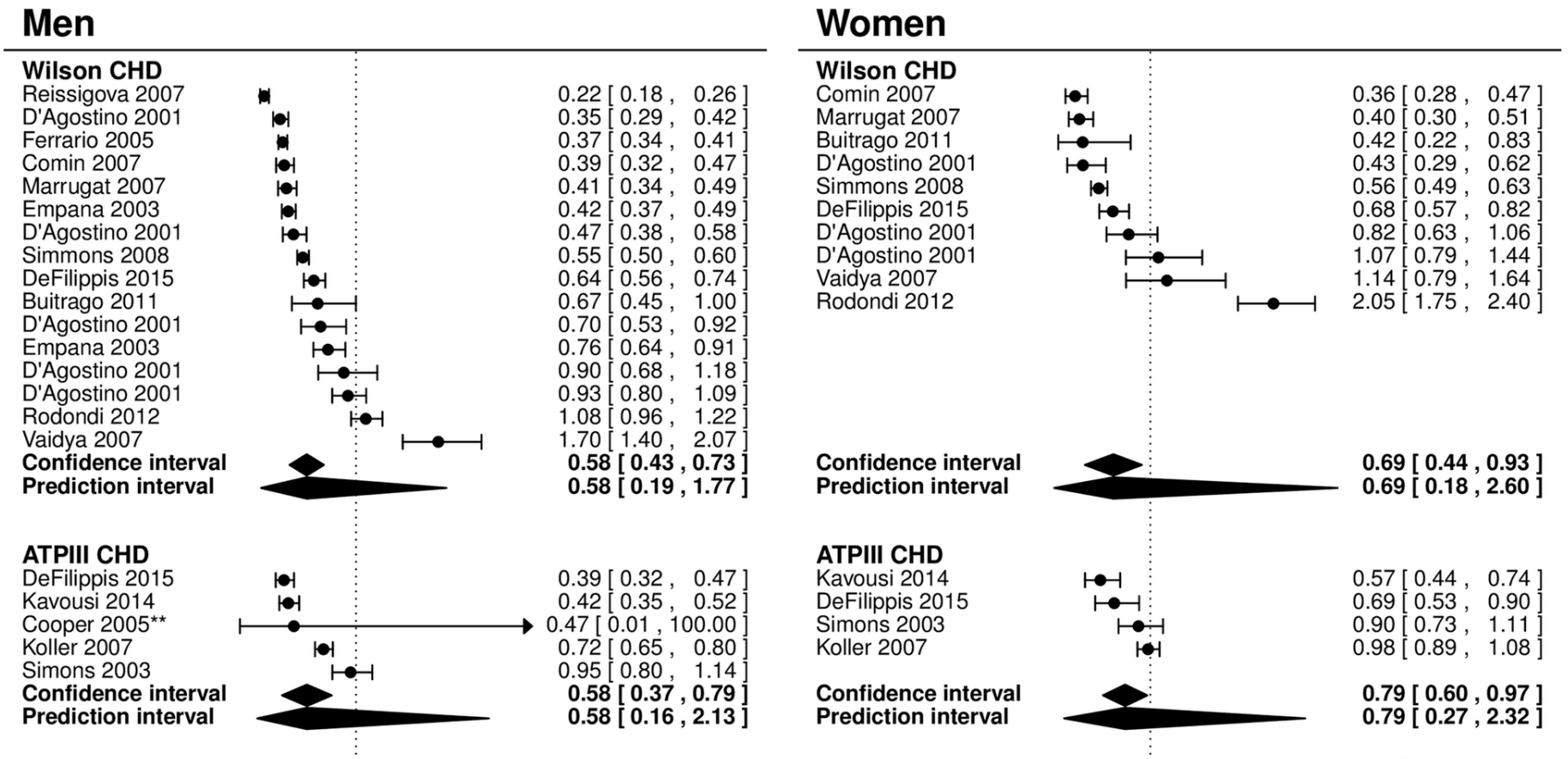
An illustrative example

Critical appraisal (PROBAST)



An illustrative example

Meta-analysis (Total O:E ratio)



An illustrative example

Meta-analysis (concordance statistic)

- Framingham Wilson
  - Men: 0.68 (95% PI: 0.61 to 0.73)
  - Women: 0.71 (95% PI: 0.51 to 0.85)
- Framingham ATP III
  - Men: 0.64 (95% PI: 0.48 to 0.77)
  - Women: 0.66 (95% PI: 0.63 to 0.69)

An illustrative example

Meta-analysis (calibration slope)

- Framingham Wilson
  - Men: 1.01 (95% PI: 0.95 to 1.07)
  - Women: 0.97 (95% PI: -0.06 to 2.00)
- Framingham ATP III
  - Men: 1.29 (95% PI: 0.14 to 2.45)
  - Women: 0.95 (95% PI: 0.87 to 1.03)

An illustrative example

Heterogeneity & interpretation

- Small differences in pooled performance (except between men and women)
- Overestimation of CHD risk (particularly in EU populations as compared to US)
- Mis-calibration appears to occur in baseline risk only
- Discrimination increases as populations become more diverse

Conclusion: Framingham models appear adequate for risk prediction, but local revisions are necessary.

What next?

Following the results of a systematic review & meta-analysis, we may decide to:

- Directly implement an existing model
  - Is any mis-calibration acceptable in terms of decision making?
- Update an existing model (e.g. Framingham Wilson)
  - Which model should be chosen? (e.g. model with best overall performance, or model with least heterogeneity in performance?)
- Develop a new model from scratch
  - Ignore prior research & sustain overfitting?
- Combine and update multiple existing models

Aggregation of prediction models

Research Article | Full Access

Meta-analysis and aggregation of multiple published prediction models

Thomas P.A. Debray, Hendrik Koffijberg, Daan Nieboer, Yvonne Vergouwe, Ewout W. Steyerberg, Karel G.M. Moons

First published: 14 January 2014 | https://doi.org/10.1002/sim.6080 | Citations: 20

Research Article | Full Access

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

Thomas P.A. Debray, Hendrik Koffijberg, Yvonne Vergouwe, Karel G.M. Moons, Ewout W. Steyerberg

First published: 26 June 2012 | https://doi.org/10.1002/sim.5412 | Citations: 22

Aggregation of prediction models

General idea

- Identify promising literature models
  - Systematic review
  - Critical appraisal
- Collect a small sample of the target population
  - Intended for validation & updating purposes
- Combine the literature models into a single model
  - The predictor–outcome associations from the original models are weighted according to their performance in the validation sample
  - The aggregated model is adjusted for the local circumstances.

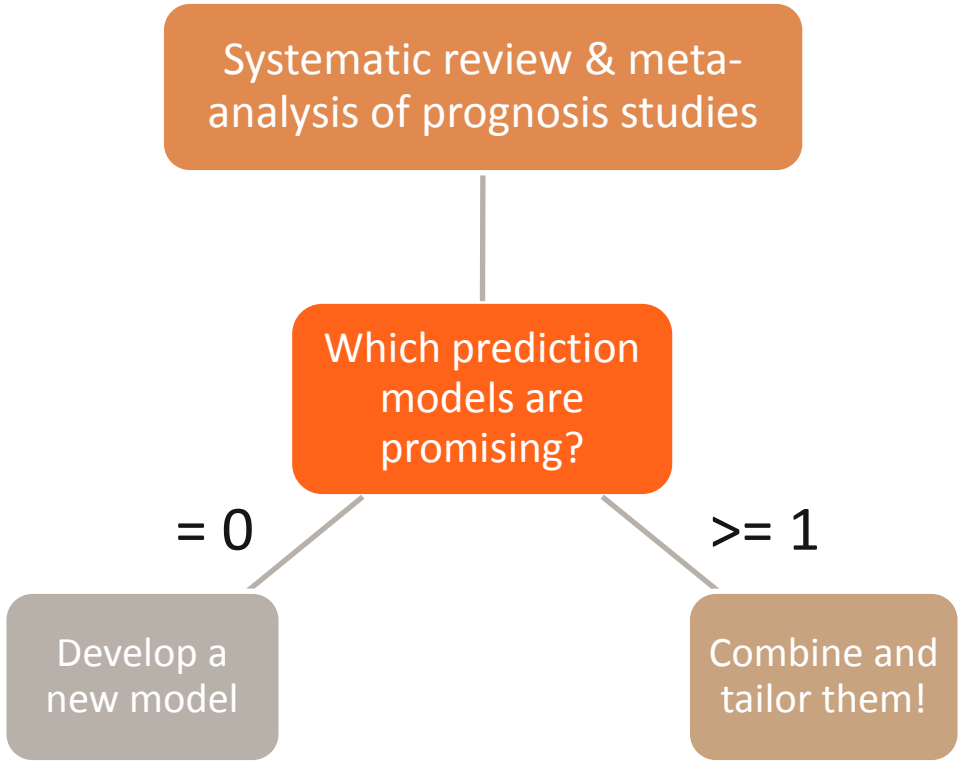
Aggregation of prediction models

Proposed approach

Stacked regressions

- Simultaneously updates, weights, and estimates the (aggregated) meta-model
- Can be viewed as a generalization of model updating
- Can be used to combine models that are poorly reported
- Effective in small samples
- Recent extensions to facilitate revision of specific predictors

The bigger picture



Ideally using IPD from multiple studies, so we can increase and assess generalizability

Ideally using a small, local sample

Summary points

- No need to develop new models
  - Systematic review and meta-analysis may help to establish whether existing models are promising
  - Identify, refine and combine promising models
  - Methods, guidance & software widely available

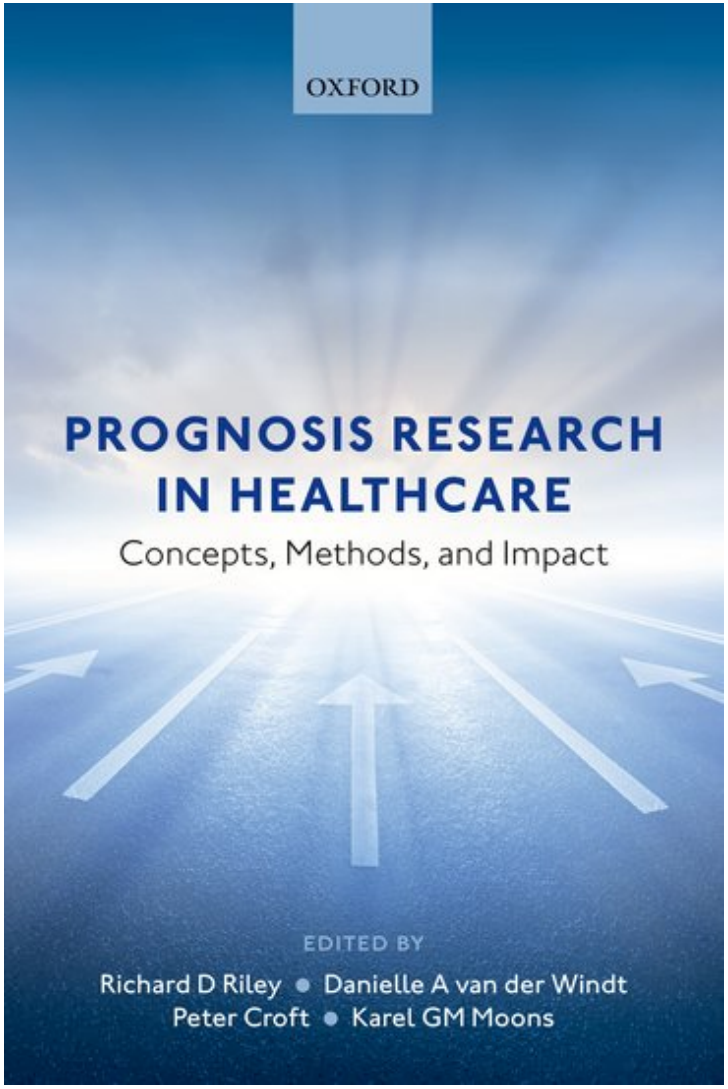
- Meta-analysis of individual participant data
  - Increase sample size and diversity in case-mix
  - Allow investigation of generalizability across different settings and populations
  - Research ongoing to address heterogeneity, missing data, measurement error, and other challenges.

Cochrane ZonMw ReCoDID

Key references

- A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis.
  Debray TP, et al. Stat Med. 2013.
- A new framework to enhance the interpretation of external validation studies of clinical prediction models.
  Debray TP, et al. J Clin Epidemiol. 2015.
- External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges.
  Riley RD, et al. BMJ. 2016.
- Construction and validation of a prognostic model across several studies, with an application in superficial bladder cancer.
  Royston P, et al. Stat Med. 2004.
- Assessment of heterogeneity in an individual participant data meta-analysis of prediction models: an overview and illustration.
  Steyerberg EW, et al. Stat Med. Under Review.

Key references



RESEARCH

Meta-analysis in prognosis research

Thomas PA Debray[1,2*†], Valentijn MT de Jong[1†], Karel GM Moons[1,2] and Richard D Riley[3]

BMC Diagnostic and Prognostic Research 2019 (Under Review)

OPEN ACCESS

GUIDELINES AND GUIDANCE

PLOS | MEDICINE

Individual Participant Data (IPD) Meta-analyses of Diagnostic and Prognostic Modeling Studies: Guidance on Their Use

Thomas P. A. Debray, Richard D. Riley, Maroeska M. Rovers, Johannes B. Reitsma, Karel G. M. Moons, Cochrane IPD Meta-analysis Methods group

Published: October 13, 2015 • https://doi.org/10.1371/journal.pmed.1001886

metamisc
Diagnostic and Prognostic Meta-Analysis

https://CRAN.R-project.org/package=metamisc

Acknowledgements











@TPA_Debray

Keele UNIVERSITY

UNIVERSITY OF OXFORD

Cochrane