Developing and validating prediction models using large clustered datasets and meta-analysis

Thomas Debray, PhD
Assistant Professor
Julius Center for Health Sciences and Primary Care

December 1, 2016
Prediction

- Risk prediction = foreseeing / foretelling
  ... (probability) of something that is yet unknown

- Turn available information (predictors) into a statement about the probability:
  ... of having a particular disease -> diagnosis
  ... of developing a particular event -> prognosis

- Use of prognostic information:
  – to inform patients and their families
  – to guide treatment and other clinical decisions
  – to create risk groups
  – ...
How do we predict?

• Combine information from multiple predictors
  – Subject characteristics (e.g. age, gender)
  – History and physical examination results (e.g. blood pressure)
  – Imaging results
  – (Bio)markers (e.g. coronary plaque)

• Develop a multivariable statistical model
  – Need for patient data from large cohort studies
  – Many strategies available (Regression, decision trees, neural networks, ...)

[Graph and image of a person diagram]
What is a good model?

**CALIBRATION**
Accurate predictions.

**GENERALIZABILITY**
Good and consistent performance across different settings and populations.

**DISCRIMINATION**
Ability to distinguish between different risk groups.

**IMPLEMENTATION**
Influence decision making.

**IMPACT**
Improve patient outcomes.
Common pitfalls

Most models are not as good as we think

- Limited sample size
- Flaws in design & analysis
- Incomplete/selective reporting
- Lack of external validation
Most models are not as good as we think

Clinical prediction models are not being validated

Peter Tugwell, J. André Knothnerus (Editors)

External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination

George C.M. Siontis\textsuperscript{a,1}, Ioanna Tzoulaki\textsuperscript{a,b}, Peter J. Castaldi\textsuperscript{c}, John P.A. Ioannidis\textsuperscript{d,e,f,*}

External validation of multivariable prediction models: a systematic review of methodological conduct and reporting

Gary S Collins\textsuperscript{1*}, Joris A de Groot\textsuperscript{2}, Susan Dutton\textsuperscript{1}, Omar Omar\textsuperscript{1}, Milensu Shanyinde\textsuperscript{1}, Abdelouahid Tajer\textsuperscript{1}, Merryn Voysey\textsuperscript{1}, Rose Wharton\textsuperscript{1}, Ly-Mee Yu\textsuperscript{1}, Karel G Moons\textsuperscript{2} and Douglas G Altman\textsuperscript{1}

Assessment of Claims of Improved Prediction Beyond the Framingham Risk Score

Ioanna Tzoulaki, PhD; George Liberopoulos, MD; John P. A. Ioannidis, MD
Most models are not as good as we think

Siontis et al (J Clin Epidemiol 2014)
Most models are not as good as we think

Lack of generalizability across settings & (sub)populations

- Poor reproducibility
  - Overfitting to the data at hand

- Poor transportability
  - Differences in patient spectrum
  - Differences in measurement methods
  - Changes in standards of care and treatment strategies
  - ...
Use of large, clustered, datasets

Research partially funded by my VENI grant (91617050) “Better predictions using big data sets”
Use of large clustered data sets

Research Methods & Reporting

External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges

BMJ 2016;353 doi:https://doi.org/10.1136/bmj.i3140 (Published 22 June 2016)
Cite this as: BMJ 2016;353:i3140

GUIDELINES AND GUIDANCE

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

Thomas P. A. Debray¹,²*, Richard D. Riley³, Maroeska M. Rovers⁴, Johannes B. Reitsma¹,², Karel G. M. Moons¹,², Cochrane IPD Meta-analysis Methods group¹
Use of large clustered data sets

Potential advantages

• Development of better prediction models
  – Reduced risk of overfitting
  – Ability to address wider spectrum of patients
  – Ability to investigate more complex associations

• More extensive testing of model performance
  – Ability to externally validate across multiple settings
  – Ability to investigate sources of heterogeneity
  – Ability to improve or tailor the model
Model development

• Need to identify whether aggregation of multiple data sources is justifiable
  – Differences in included populations
  – Differences in measurement methods
  – Differences in treatment standards

• Need to account for heterogeneity across settings
  – Differences in outcome prevalence (or incidence)
  – Differences in predictor effects
  – Failing to account for clustering may cause prediction models to yield poor performance across different (sub)populations!

Implement a framework that uses internal-external cross-validation
A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis

Thomas P.A. Debray, Karel G.M. Moons, Ikhlaaq Ahmed, Hendrik Koffijberg, Richard David Riley

First published: 11 January 2013  Full publication history

Commentary

Prediction models need appropriate internal, internal–external, and external validation

Ewout W. Steyerberg, Frank E. Harrell Jr.
Internal-external cross-validation (IECV)

Pre-defined development strategy

V → IPD-1 → Performance study 1
D → IPD-2 → Performance study 2
D → IPD-3 → Performance study 3
D → IPD-4 → Performance study 4

Summarize estimates of model performance and assess model generalizability

Develop final model
Internal-external cross-validation (IECV)

The IECV approach allows for many external validations
Assessing model performance

Synthesis of performance estimates

• A ‘good’ prediction model should have
  – satisfactory performance on average
  – little or no between-study heterogeneity in performance

• Meta-analysis may help ...
  – To estimate likely performance in new studies
  – To identify sources of heterogeneity
  – To evaluate different modeling strategies
  – To distinguish between reproducibility and transportability
  – To identify “boundaries” of model generalizability
Assessing model performance

A new framework to enhance the interpretation of external validation studies of clinical prediction models

Thomas P.A. Debray\textsuperscript{a,*}, Yvonne Vergouwe\textsuperscript{b}, Hendrik Koffijberg\textsuperscript{a}, Daan Nieboer\textsuperscript{b}, Ewout W. Steyerberg\textsuperscript{b,1}, Karel G.M. Moons\textsuperscript{a,1}

Multivariate meta-analysis of individual participant data helped externally validate the performance and implementation of a prediction model

Kym I.E. Snell\textsuperscript{a}, Harry Hua\textsuperscript{b}, Thomas P.A. Debray\textsuperscript{c,d}, Joie Ensor\textsuperscript{e}, Maxime P. Loom\textsuperscript{f}, Karel G.M. Moons\textsuperscript{c,d}, Richard D. Riley\textsuperscript{e,*}

Meta-analysis of prediction model performance across multiple studies: Which scale helps ensure between-study normality for the C-statistic and calibration measures?

Kym I.E. Snell, Joie Ensor, Thomas PA Debray,...

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

Thomas P A Debray,\textsuperscript{1,2} Johanna A A G Damen,\textsuperscript{1,2} Kym I E Snell,\textsuperscript{3} Joie Ensor,\textsuperscript{3} Lotty Hooft,\textsuperscript{1,2} Johannes B Reitsma,\textsuperscript{1,2} Richard D Riley,\textsuperscript{3} Karel G M Moons\textsuperscript{1,2}
Assessing model performance

Investigation of heterogeneity across settings

Concordance statistic for QRISK2, for each of the 364 included general practices (N = 2,000,000)

Calibration of QRISK2 and the Framingham risk score in women aged 35 to 74 years

Summary C statistic = 0.83 (95% CI 0.826 to 0.833)
95% prediction interval for true C statistic in a new practice = 0.76 to 0.88

More guidance underway!

- Future book chapters
  - Collins GS, Moons KGM, Debray TPA, Altman DG, Riley RD. *Systematic reviews of prediction models*. Systematic Reviews in Health Care: Meta-Analysis in Context (Wiley)
  - Riley TD, Debray TPA, Moons KGM. *Individual Participant Data Meta-analysis of Prognosis Studies*. Evidence synthesis using individual participant data: concepts, methods, and guidance for clinical research (CRC Press)

- Courses
  - MSc Epidemiology
  - Elevate
  - Cochrane