UMC Utrecht

Recent advances in prediction modelling methods

Thomas Debray, PhD
Assistant Professor

@TPA_Debray

10 Aug 2018, Sydney

University Medical Center Utrecht

UMC Utrecht

Risk 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**

Why do we predict?

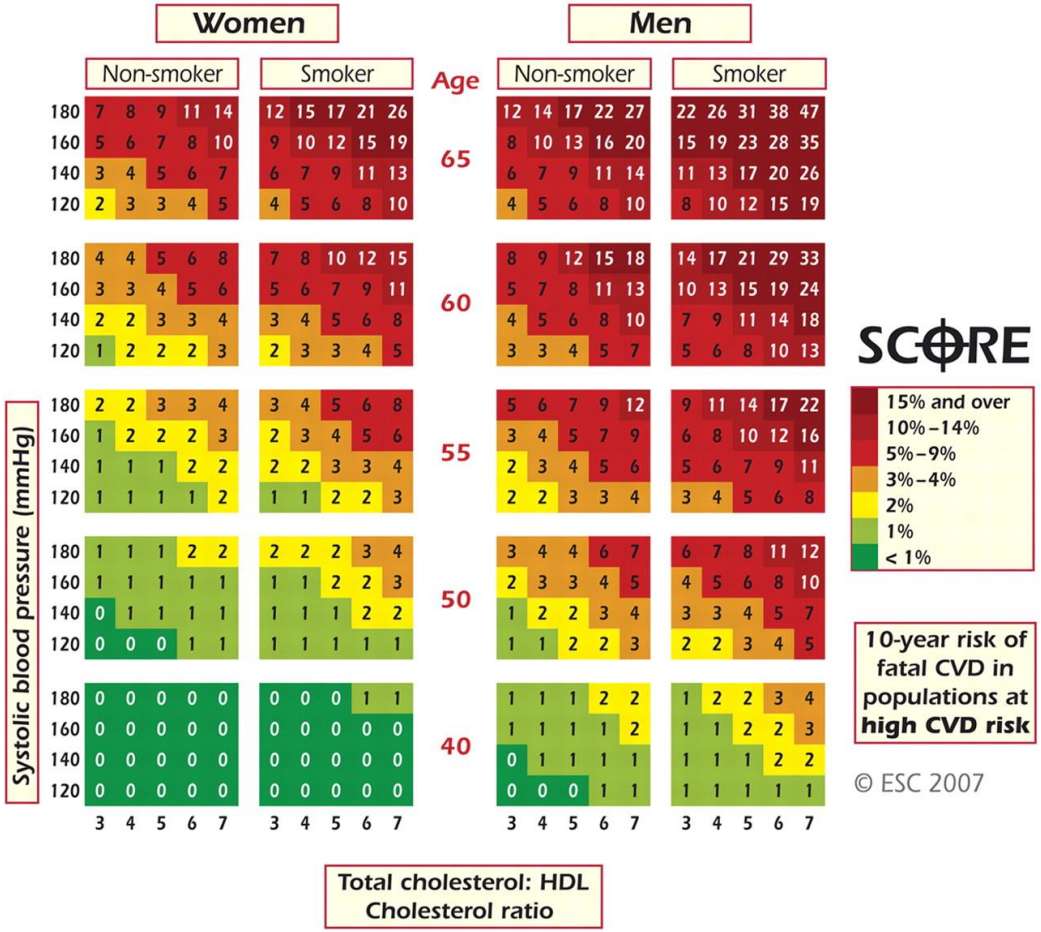
- Identification of high risk individuals
  - To inform patients and their families
  - To guide treatment decisions ("**precision medicine**")
  - To design randomized trials

- Data analysis
  - To deal with missing values
  - To match/subclassifiy patients
  - …

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 individual participant data (e.g. from cohort studies)
  - Many strategies available (e.g. logistic regression)


Women
Men
Non-smoker
Smoker
Non-smoker
Smoker
Age
Systolic blood pressure (mmHg)
180 7 8 9 11 14 12 15 17 21 26 12 14 17 22 27 22 26 31 38 47
160 5 6 7 8 10 9 10 12 15 19 8 10 13 16 20 15 19 23 28 35
140 3 4 5 6 7 6 7 9 11 13 6 7 9 11 14 11 13 17 20 26
120 2 3 3 4 5 4 5 6 8 10 4 5 6 8 10 8 10 12 15 19
65
180 4 4 5 6 8 7 8 10 12 15 8 9 12 15 18 14 17 21 29 33
160 3 3 4 5 6 5 6 7 9 11 5 7 8 11 13 10 13 15 19 24
140 2 2 3 3 4 3 4 5 6 8 4 5 6 8 10 7 9 11 14 18
120 1 2 2 2 3 2 3 3 4 5 3 3 4 5 7 5 6 8 10 13
60
180 2 2 3 3 4 3 4 5 6 8 5 6 7 9 12 9 11 14 17 22
160 1 2 2 2 3 2 3 4 5 6 3 4 5 7 9 6 8 10 12 16
140 1 1 1 2 2 2 2 3 3 4 2 3 4 5 6 5 6 7 9 11
120 1 1 1 1 2 1 1 2 2 3 2 2 3 3 4 3 4 5 6 8
55
180 1 1 1 2 2 2 2 2 3 4 3 4 4 6 7 6 7 8 11 12
160 1 1 1 1 1 1 1 2 2 3 2 3 3 4 5 4 5 6 8 10
140 0 1 1 1 1 1 1 1 2 2 1 2 2 3 4 3 3 4 5 7
120 0 0 0 1 1 1 1 1 1 1 1 1 2 2 3 2 2 3 4 5
50
180 0 0 0 0 0 0 0 0 1 1 1 1 2 2 1 2 2 3 4
160 0 0 0 0 0 0 0 0 0 0 1 1 1 1 2 1 1 2 2 3
140 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 2 2
120 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1
40
3 4 5 6 7 3 4 5 6 7 3 4 5 6 7 3 4 5 6 7
Total cholesterol: HDL Cholesterol ratio
SCORE
15% and over
10%–14%
5%–9%
3%–4%
2%
1%
< 1%
10-year risk of fatal CVD in populations at high CVD risk
© ESC 2007

@TPA_Debray

IBM Watson


PIENSE
THINK
ΣΚΕΨΟΥ
DENKE
PENSER
$300,000
Who is Stoker?
(I FOR ONE WELCOME OUR NEW COMPUTER OVERLORDS)
$1,000
$1,000,000
Who is Bram Stoker?
$ 17,973
$200,000
WHO IS BRAM STOKER?
$5600

Watson for Oncology

"Bring personalized, evidence-supported cancer care plans to your patients"

- Interpret cancer patients' clinical information
- Digest doctor's notes, medical studies, and clinical guidelines
- Provide individualized treatment recommendations
- Adopted by more than 150 hospitals and healthcare organizations across 11 countries, including China

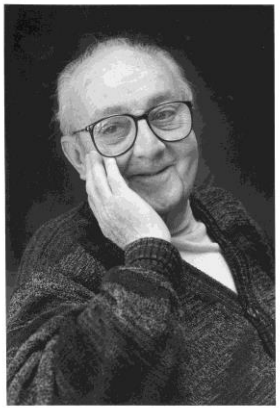
https://www.ibm.com/watson/health/oncology-and-genomics/oncology/

Hype meets reality

- Focus on US clinical practice and demographics
- Reliance on varies among hospitals
- Multiple examples of unsafe and incorrect treatment recommendations
- Lack of validation by independent scientists
- Lack of clinical trials to assess effectiveness

https://www.statnews.com/2017/09/05/watson-ibm-cancer/

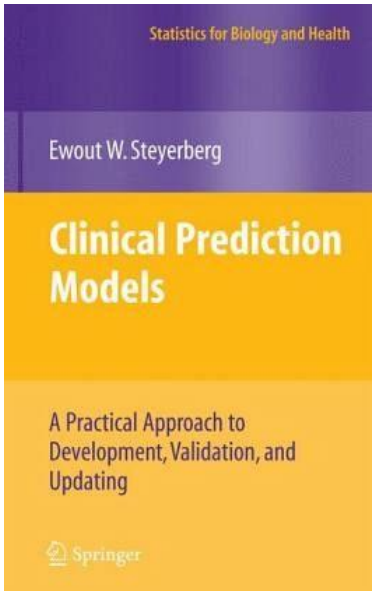
Most models are not as good as we think

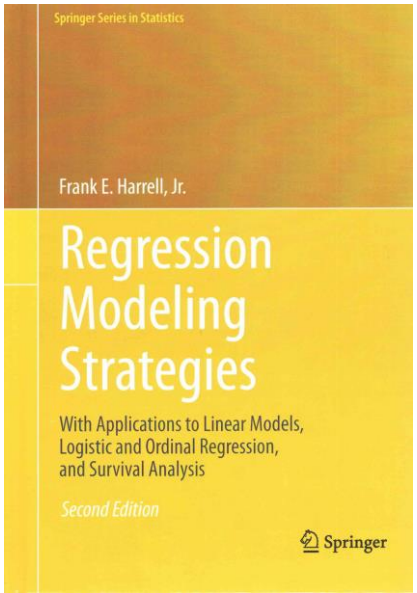


“All models are wrong, but some are useful”

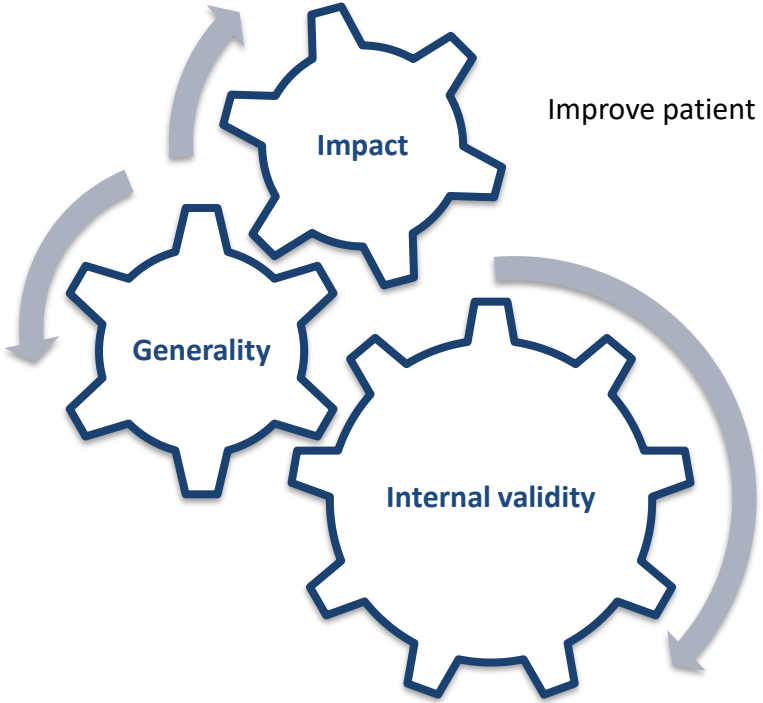
George Box

What is a “good” prediction model?





What is a “good” prediction model?



Improve patient outcomes

Good and consistent performance across different settings and populations

Discrimination
Ability to distiguish between low risk and high risk patients

Calibration
Accurate risk predictions

(A selection of) recent advances

- Evidence Synthesis
- Big data
- Modeling of treatment
- Guidance and software

Evidence synthesis

UMC Utrecht

Evidence synthesis

Integrating published evidence during model development

- Prognostic factor studies (e.g. factor-outcome associations)
- Prognostic model studies (e.g. prediction models)


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]

Submitted to BMC Diagnostic and Prognostic Research

Synthesis of prognostic model studies

Numerous published models for same target population and outcomes

- > 300 models alike Framingham, SCORE, QRisk
- > 100 models for brain trauma patients
- > 100 diabetes type 2 models
- > 60 models for breast cancer prognosis

Synthesis of prognostic model studies

Combine and tailor previously published models

- Two-stage meta-analysis of regression coefficients
- Mixture modeling
- Model averaging
- Stacked regressions
- Principal components regression
- Multivariate generalized least squares

Synthesis of prognostic model studies

Combine and tailor previously published models

- Two-stage meta-analysis of regression coefficients
- Mixture modeling
- Model averaging
- Stacked regressions
- Principal components regression
- Multivariate generalized least squares

Model averaging

1. Validate and update existing models in “own” IPD
2. Calculate a model-specific prediction for each patient
3. Average model predictions for each patient
   - Assign more weight to models with better fit in the IPD
   - Assign less weight to models that have been substantially revised

$$w_m = \frac{\exp(-0.5\,\mathrm{BIC}_m)}{\sum_{l=1}^{M} \exp(-0.5\,\mathrm{BIC}_l)}$$

4. Use the models’ averaged predictions as dependent variable to develop the meta-model

Stacked regressions

1. Treat the predictions of each literature model as independent variable
2. Generate a linear combination of the model predictions
   - Estimation of a common intercept term
   - Estimation of a regression coefficient for each model
   - Omit models with a “negative” contribution
3. Calculate regression coefficients of the meta-model by applying the estimated weights

Simultaneous updating, discovery and estimation of the best combination of literature models

Diagnosis of Deep Vein Thrombosis

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)).

Clinical feature	Score
Active cancer (treatment ongoing or within previous 6 months or palliative)	1
Paralysis, paresis, or recent plaster immobilisation of the lower extremities	1
Recently bedridden for more than 3 days or major surgery, within 4 weeks	1
Localised tenderness along the distribution of the deep venous system	1
Entire leg swollen	1
Calf swelling by more than 3 cm when compared with the asymptomatic leg (measured 10 cm below tibial tuberosity)	1
Pitting oedema (greater in the symptomatic leg)	1
Collateral superficial veins (non-varicose)	1
Alternative diagnosis as likely or greater than that of deep-vein thrombosis	−2

In patients with symptoms in both legs, the more symptomatic leg is used.

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

Diagnosis of Deep Vein Thrombosis

- Previously published prediction models for diagnosing DVT
  - Wells
  - Modified Wells
  - Gagne
  - Hamilton
  - Oudega

- Patient-level data
  - Primary Care dataset (N=1028)
  - We applied stacked regressions to combine the models

Diagnosis of Deep Vein Thrombosis

	Weight	X 1	X 2	X 3	X 4	X 5	X 6	X 7	X 8	X 9	X 10	X 11	X 12	X 13	X 14	X 15
Wells [2]	0	■	■	■	■	■	■	■	■	■						
Modified Wells [2]	0	■	■	■	■	■	■	■	■	■	■					
Gagne [1]	0.497	■		■			■			■	■	■				
Hamilton [2]	0	■	■	■			■				■		■			
Oudega [1]	0.537	■		■			■		■			■	■	■	■	■
Stacked Regressions		■		■			■		■	■	■	■	■	■	■	■

Final model: $\Pr(\text{DVT present}) = \text{logit}^{-1}(-3.6 + 1.2\, x_{\text{malign}} + 0.5\, x_{\text{surg}} + x_{\text{cdif3}} + 0.3\, x_{\text{vein}} - x_{\text{adiag}} + 0.4\, x_{\text{histdvt}} + x_{\text{oachst}} + 0.3\, x_{\text{sex}} + 0.3\, x_{\text{notraum}} + 1.6\, x_{\text{ddimd}})$

Only 6 (rather than 12) degrees of freedom were needed for estimation

Big data

UMC Utrecht

The rise of “big” data sets

The rise of “big” data sets

Data increasingly available for thousands or even millions of patients from multiple practices, hospitals, or countries.

- Meta-analysis of individual participant data from multiple studies
  - Observational studies
  - Randomized controlled trials

- Analyses of databases and registry data containing e-health records

Examples of “big” data sets

International Prediction of Pre-eclampsia IPD Collaborative Network

- Target population
  - Pregnant women in the 1st or 2nd trimester of pregnancy

- So far, 81 datasets have been included
  - 15 UK studies
  - 66 international studies

Examples of “big” data sets

CALIBER

- EHR data encompassing more than 10 million adults with 400 million person-years of follow-up
- Primary care consultations and hospitalisations
- Clinical examination findings, blood laboratory results, prescriptions and vaccinations
- Diagnoses of diseases and mortality data

Why do we need “big”data sets?

- Development of better prediction models
  - Reduced risk of overfitting
  - Ability to address wider spectrum of patients
  - Ability to estimate more complex associations

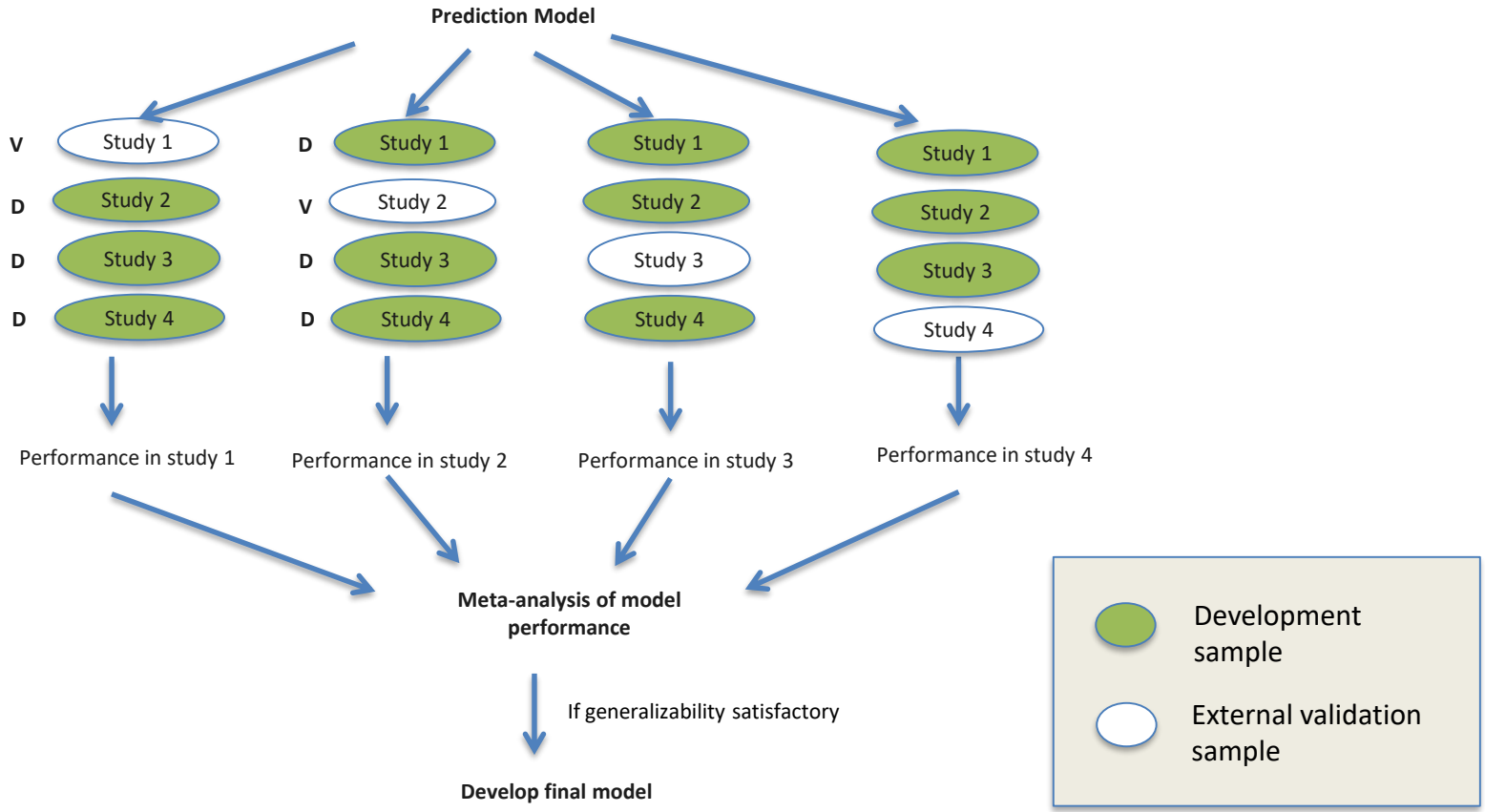
- More extensive testing of model performance, as to establish whether model performance is
  - Satisfactory on average
  - Consistently good across different settings and (sub)populations

10.1371/journal.pmed.1001886

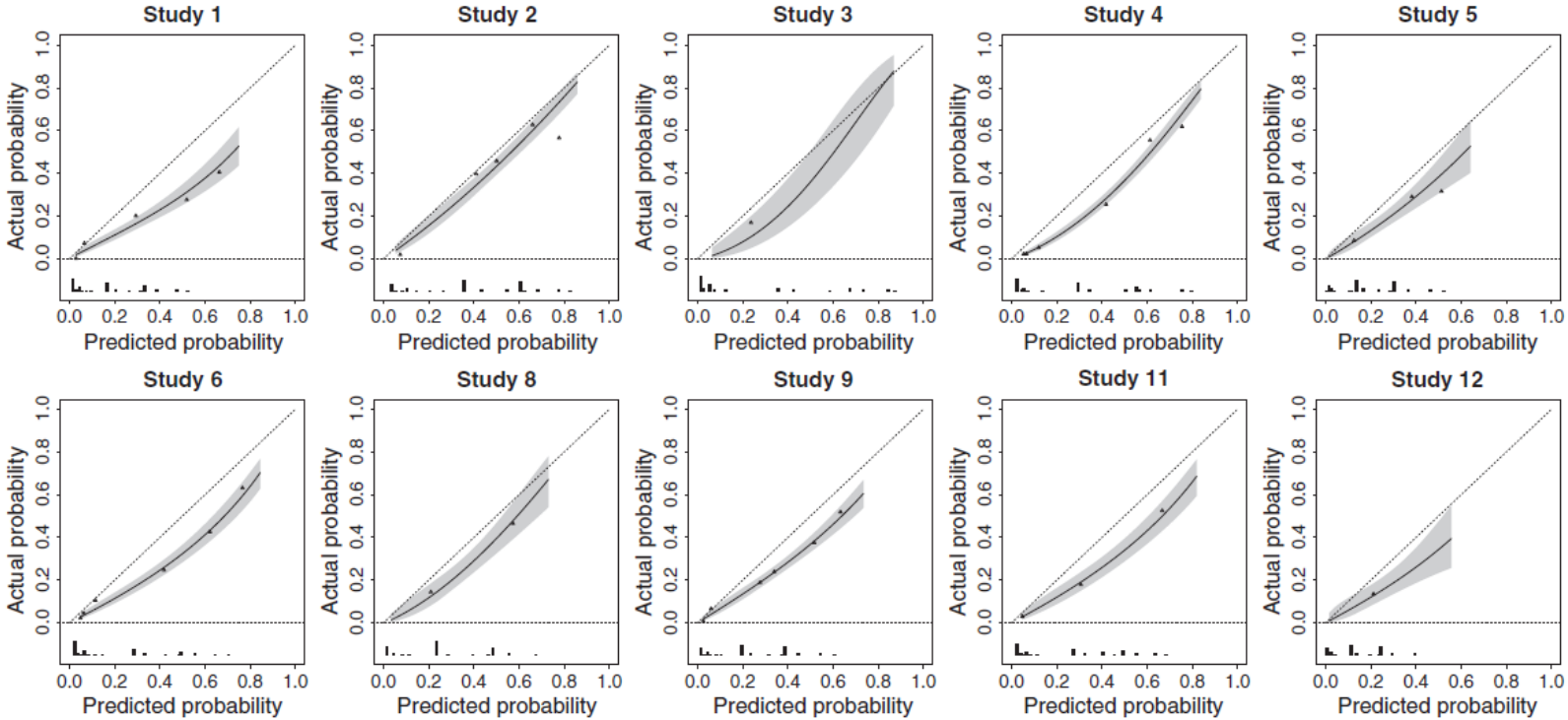
Internal-external cross-validation

- Royston *et al*. Construction and validation of a prognostic model across several studies, with an application in superficial bladder cancer. Stat Med 2004.
- Debray *et al*. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. Stat Med 2013.
- Steyerberg and Harrell. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol 2015.

Internal-external cross-validation (IECV)

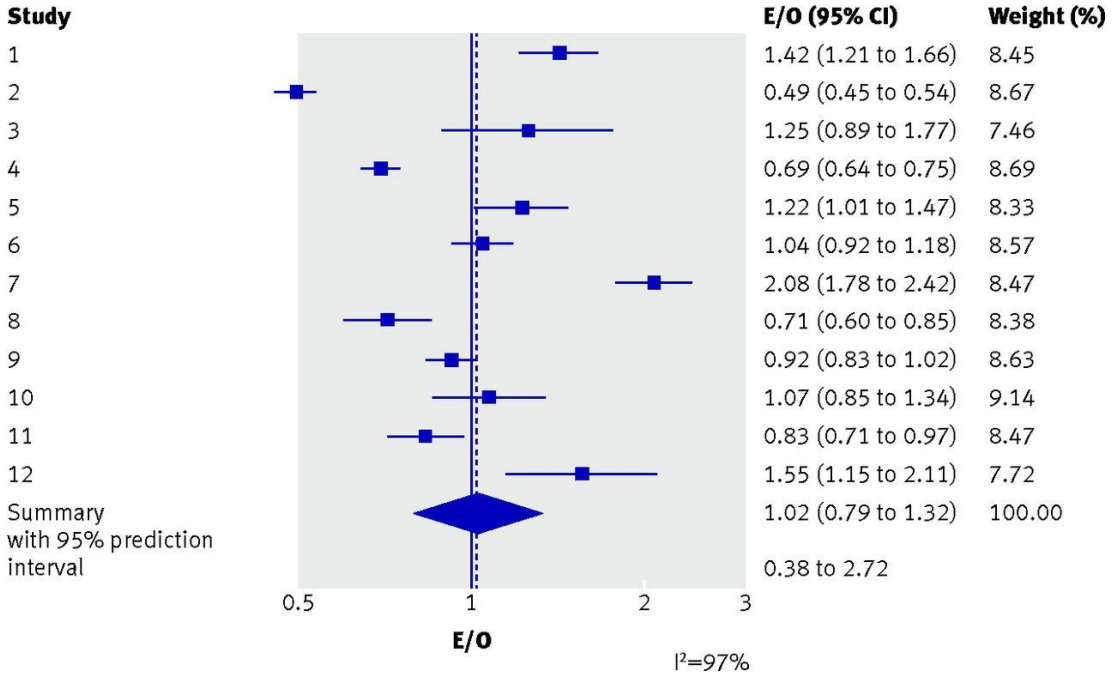


IECV allows for many external validations



10.1002/sim.5732

IECV allows for meta-analysis



10.1136/bmj.i3140

Development and validation of ENCALS

Prognosis for patients with amyotrophic lateral sclerosis (ALS)

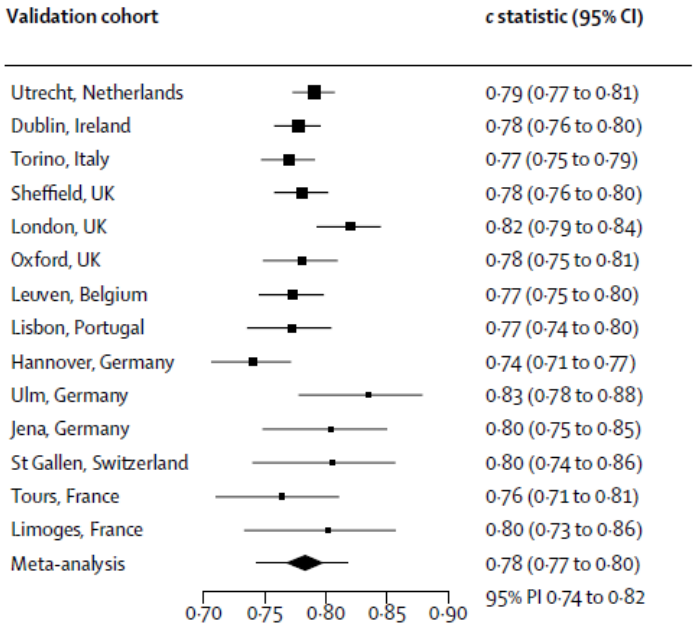
- Cohort data from 11,475 patients from 14 European ALS centres
- Composite survival outcome (non-invasive ventilation for more than 23 h per day, tracheostomy, or death)
- Development of multivariable Royston-Parmar models
- Assessment of generalizability via IECV

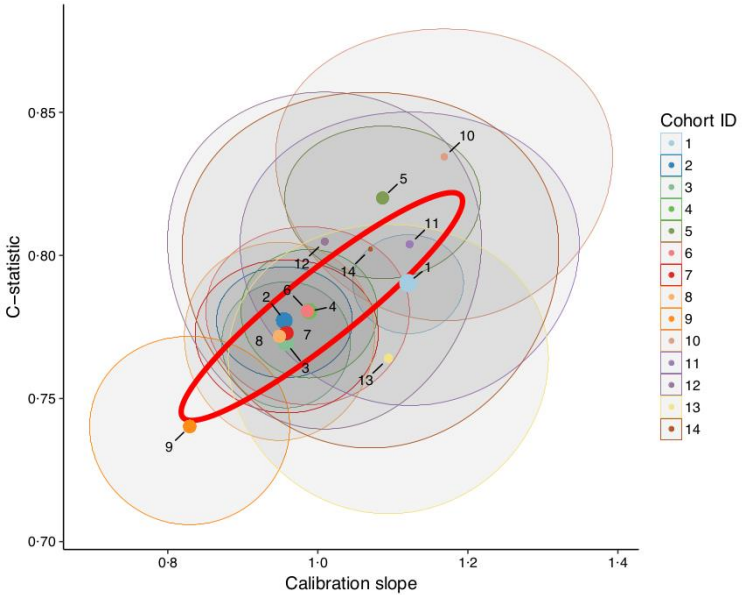
THE LANCET Neurology

@TPA_Debray

10.1016/S1474-4422(18)30089-9

Development and validation of ENCALS





@TPA_Debray

10.1016/S1474-4422(18)30089-9

Development and validation of ENCALS

Measure	Criteria	Prob. of "good" performance	Joint probability
C-statistic	> 0.70	100%	98.3%
Calibration slope	0.80 to 1.20	97.1%	
Calibration-in-the-large	-0.587 to 0.587	85.5%	

10.1016/S1474-4422(18)30089-9

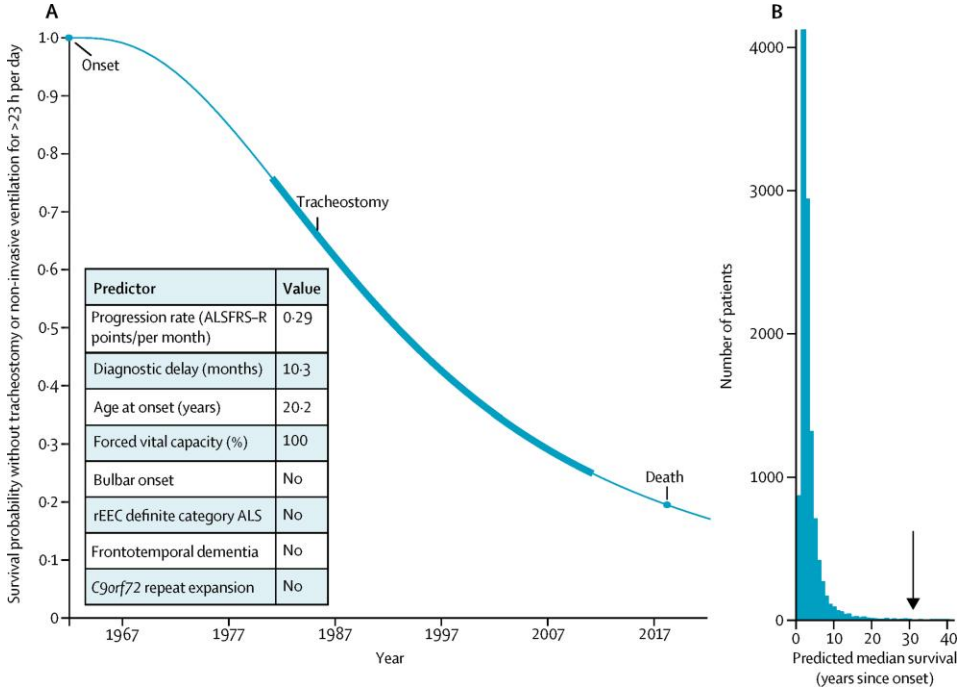
The life expectancy of Stephen Hawking

"Using publicly available data, we examined whether Professor Hawking's survival was as rare as his intellectual performance, or could be predicted solely based on his disease characteristics at diagnosis in 1963."

- Predicted 10-year survival probability: 94%
- The IQR for his predicted survival lay between 1981 and 2011
- Young age of onset was the most important factor for his long survival

10.1016/S1474-4422(18)30241-2

The life expectancy of Stephen Hawking



10.1016/S1474-4422(18)30241-2

Modeling of treatment

UMC Utrecht

The role of treatment

- Many treatments are a strong prognostic factor

- Treatment efficacy may
  - change over time
  - vary across settings and populations

- Prediction models are usually developed …
  - In untreated individuals
  - In an arbitrary mixture of treated and untreated individuals
  - Sometimes, a treatment indicator is included

The role of treatment

Treatment is a common source of inaccurate risk predictions

- Failure to account for received treatments
- Failure to account for treatmens started during study period
- Failure to account for confounding of treatment effects
- Failure to account for modifiers of treatment effect

Treatment should explicitly be modeled to improve the generalizability of prediction models, and to facilitate the estimation of individual response to treatment.

Recent proposals

- Machine Learning
  - Generalised Linear Model Trees with Global Additive Effects
  - Ensemble of of survival trees
  - Deep learning

- Full modelling
  - Linear effects for treatment and interactions (Groenwold, JCE 2016)
  - Penalized effects for treatment and interactions (Van Klaveren et al., JCE 2015)
  - Nonlinear interaction terms (Royston & Sauerbrei, Stat Med 2013)
  - Time-varying confounding (to adjust for treatment drop-in)

More work to do

- Dealing with multiple datasets and missing data
  - Debray et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. Res Synth Methods 2015.
  - Audigier et al. Multiple imputation for multilevel data with continuous and binary variables. Stat Sci 2018.
- Dealing with randomized and non-randomized studies
  - Efthimiou et al. Combining randomized and non-randomized evidence in network meta-analysis. Stat Med 2017.
  - Verde & Ohmann. Combining randomized and non-randomized evidence in clinical research: a review of methods and applications. Res Synth Methods 2015.
- Dealing with multiple treatments
  - Debray et al. An overview of methods for network meta-analysis using individual participant data: wh do benefits arise? Stat Methods Med Res 2018.

Guidance and software

UMC Utrecht

Guidance

- **Prognostic Research in Health Care: concepts, methods and impact**
  editors: Richard Riley, Danielle Van der Windt, Peter Croft, Karel Moons

- **Evidence synthesis using individual participant data: Concepts, Methods and Guidance for Clinical Research**
  editors: Richard Riley, Jayne Tierney, Lesley Stewart

- **Handbook of Meta-analysis**
  editors: Christopher Schmid, Theo Stijnen, Ian White

Software

metamisc: Diagnostic and Prognostic Meta-Analysis

Meta-analysis of diagnostic and prognostic modeling studies. Summarize estimates of prognostic factors, diagnostic test accuracy and prediction model performance. Validate, update and combine published prediction models. Develop new prediction models with data from multiple studies.

Version:	0.1.9
Depends:	R (≥ 3.2.0), stats, graphics
Imports:	metafor (≥ 2.0.0), mvtnorm, ellipse, lme4, plyr, ggplot2
Suggests:	runjags, rjags, testthat (≥ 1.0.2)
Published:	2018-05-13
Author:	Thomas Debray [aut, cre], Valentijn de Jong [aut]
Maintainer:	Thomas Debray <thomas.debray at gmail.com>
License:	GPL-3
URL:	http://r-forge.r-project.org/projects/metamisc/
NeedsCompilation:	no
In views:	MetaAnalysis
CRAN checks:	metamisc results

Downloads:

Reference manual:	metamisc.pdf
Package source:	metamisc_0.1.9.tar.gz
Windows binaries:	r-devel: metamisc_0.1.9.zip, r-release: metamisc_0.1.9.zip, r-oldrel: metamisc_0.1.9.zip
OS X binaries:	r-release: metamisc_0.1.9.tgz, r-oldrel: metamisc_0.1.8.tgz
Old sources:	metamisc archive

Linking:

Please use the canonical form https://CRAN.R-project.org/package=metamisc to link to this page.