Regression and Clinical prediction models

Session 3
Steps in planning and conducting CPM research – Part 2

Pedro E A A do Brasil
pedro.brasil@fiocruz.br
2023

Objectives

- Clinical prediction models have some unique characteristics which make them different from other observational studies.
- In this session, usual steps in planning and conducting CPM research will be introduced and commented.
- One must be well aware of which state of development the research line is, to know what additional evidence is necessary to have a prediction model available.

Statistical model

- Choosing the model (tool) is not an easy task
  - Many options with modern modeling and software availability
  - Often advantages of a model over another is theoretical but not confirmed on predictions accuracy
  - Medical readers may be resistant to unusual models, even if they predict better
  - Common outcomes formats helps to choose a model
    - binary, unordered categorical, ordered categorical, continuous, and survival data.

Statistical model

- Some options according to outcomes formats (e.g.)
  - For binary outcome
    - Logistic regression, decision trees, neural network, GAM, MARS, GEE, SVM, Random forest
  - Unordered categorical outcome
    - Multinomial regression, neural network
  - Ordered categorical outcome
    - Ordered logistic regression
  - Continuous outcome
    - Ordinary least squared (linear regression), GAM, SVM, GEE , neural networks
  - Survival outcome
    - Cox or parametric survival models, decision trees, neural networks, Random forest

Statistical model

- Before definitely choosing a model one may consider
  - Wonder and possibly test if model assumptions can be met only to the extent that adaptations to the model lead to better predictions
  - Wonder if model assumptions can be flexiblelized or worked around
  - Significant violations of underlying assumptions do not mean that a model predicts poorly
  - Robustness is preferred over flexibility in capturing idiosyncracies
  - Test two or more options of models
  - Transform the outcome of interest
    - To follow model assumptions or facilitate modeling and predictions
    - Be very very careful in back transforming
  - Results of the model should be transparent and presentable to the intended audience.

Statistical model

- Quality of predictions may depend on:
  - The essential quality and appropriateness of the method
  - The actual implementation of the method as a computer program
  - The skill of the “data pilot”

Statistical model

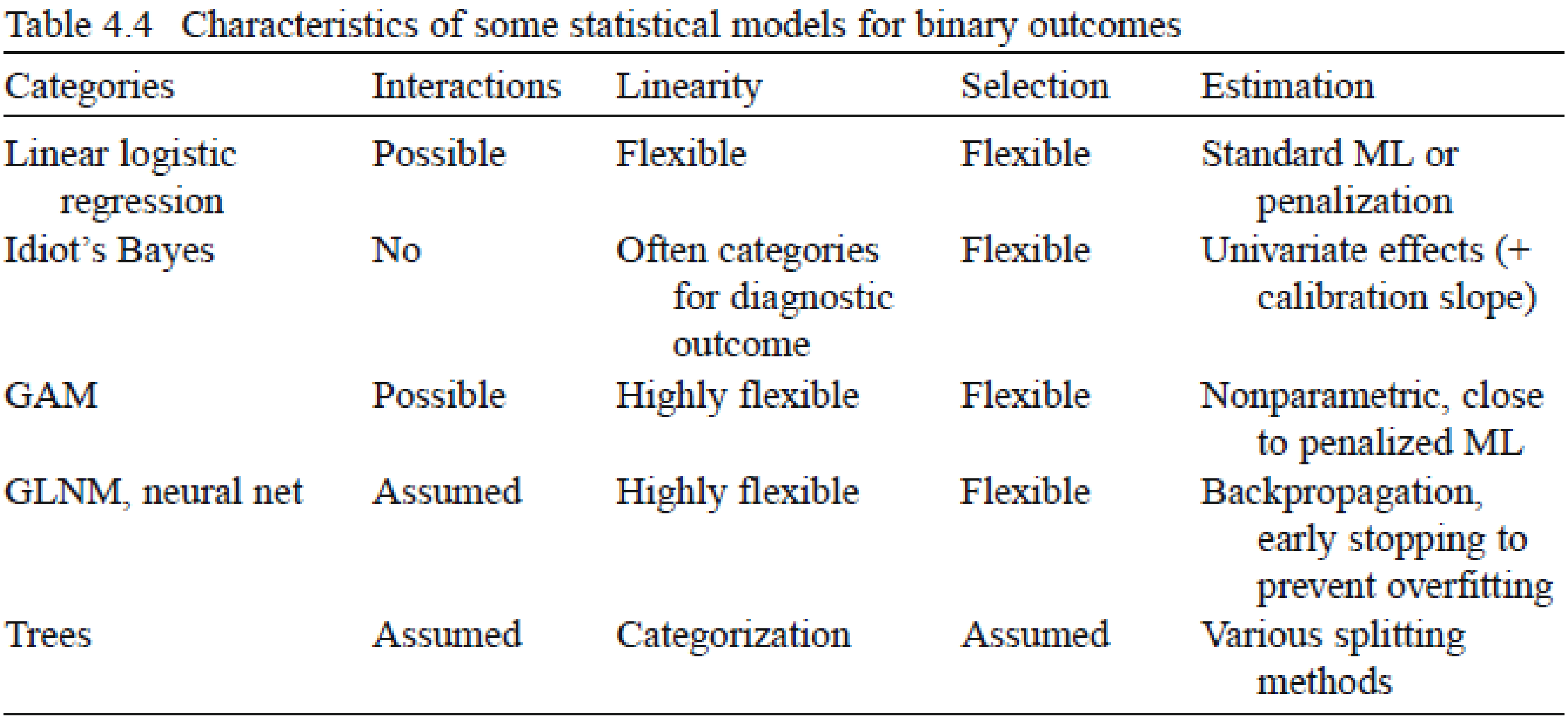


Table 4.4 Characteristics of some statistical models for binary outcomes

Categories	Interactions	Linearity	Selection	Estimation
Linear logistic regression	Possible	Flexible	Flexible	Standard ML or penalization
Idiot's Bayes	No	Often categories for diagnostic outcome	Flexible	Univariate effects (+ calibration slope)
GAM	Possible	Highly flexible	Flexible	Nonparametric, close to penalized ML
GLNM, neural net	Assumed	Highly flexible	Flexible	Backpropagation, early stopping to prevent overfitting
Trees	Assumed	Categorization	Assumed	Various splitting methods

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Statistical model

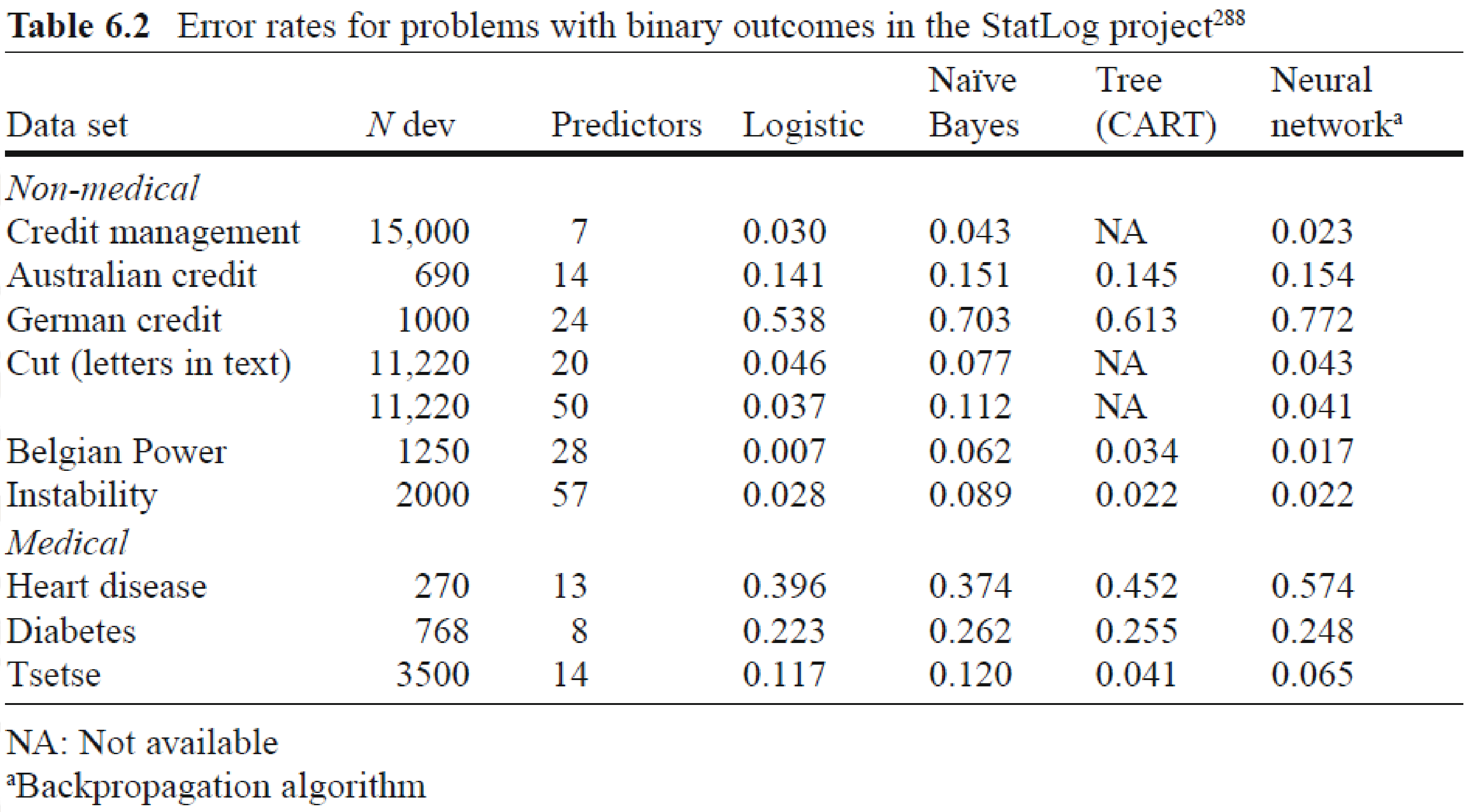


Table 6.2 Error rates for problems with binary outcomes in the StatLog project[288]

Data set	*N* dev	Predictors	Logistic	Naïve Bayes	Tree (CART)	Neural network[a]
Non-medical						
Credit management	15,000	7	0.030	0.043	NA	0.023
Australian credit	690	14	0.141	0.151	0.145	0.154
German credit	1000	24	0.538	0.703	0.613	0.772
Cut (letters in text)	11,220	20	0.046	0.077	NA	0.043
	11,220	50	0.037	0.112	NA	0.041
Belgian Power	1250	28	0.007	0.062	0.034	0.017
Instability	2000	57	0.028	0.089	0.022	0.022
Medical						
Heart disease	270	13	0.396	0.374	0.452	0.574
Diabetes	768	8	0.223	0.262	0.255	0.248
Tsetse	3500	14	0.117	0.120	0.041	0.065

NA: Not available

[a] Backpropagation algorithm

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Statistical model

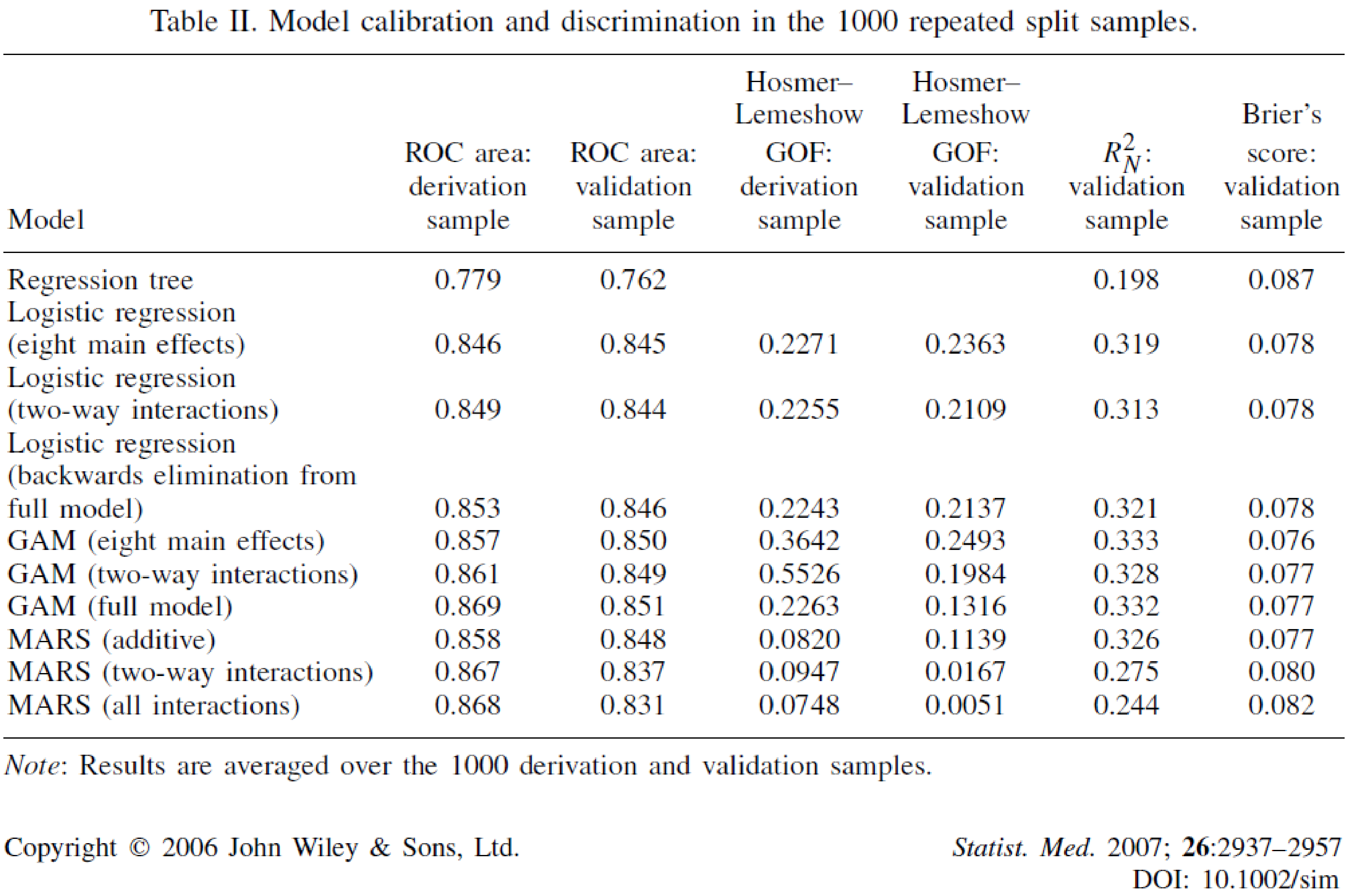


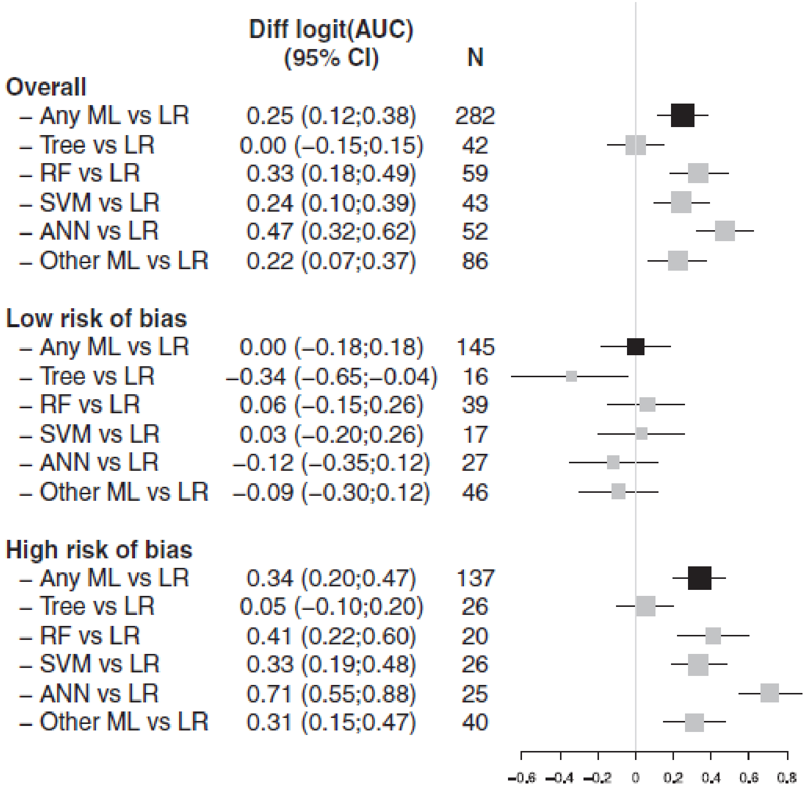
Table II. Model calibration and discrimination in the 1000 repeated split samples.

Model	ROC area: derivation sample	ROC area: validation sample	Hosmer–Lemeshow GOF: derivation sample	Hosmer–Lemeshow GOF: validation sample	R_N^2: validation sample	Brier's score: validation sample
Regression tree	0.779	0.762			0.198	0.087
Logistic regression (eight main effects)	0.846	0.845	0.2271	0.2363	0.319	0.078
Logistic regression (two-way interactions)	0.849	0.844	0.2255	0.2109	0.313	0.078
Logistic regression (backwards elimination from full model)	0.853	0.846	0.2243	0.2137	0.321	0.078
GAM (eight main effects)	0.857	0.850	0.3642	0.2493	0.333	0.076
GAM (two-way interactions)	0.861	0.849	0.5526	0.1984	0.328	0.077
GAM (full model)	0.869	0.851	0.2263	0.1316	0.332	0.077
MARS (additive)	0.858	0.848	0.0820	0.1139	0.326	0.077
MARS (two-way interactions)	0.867	0.837	0.0947	0.0167	0.275	0.080
MARS (all interactions)	0.868	0.831	0.0748	0.0051	0.244	0.082

Note: Results are averaged over the 1000 derivation and validation samples.

Statist. Med. 2007; **26**:2937–2957
DOI: 10.1002/sim

Statistical model



	Diff logit(AUC) (95% CI)	N
Overall		
– Any ML vs LR	0.25 (0.12;0.38)	282
– Tree vs LR	0.00 (–0.15;0.15)	42
– RF vs LR	0.33 (0.18;0.49)	59
– SVM vs LR	0.24 (0.10;0.39)	43
– ANN vs LR	0.47 (0.32;0.62)	52
– Other ML vs LR	0.22 (0.07;0.37)	86
Low risk of bias		
– Any ML vs LR	0.00 (–0.18;0.18)	145
– Tree vs LR	–0.34 (–0.65;–0.04)	16
– RF vs LR	0.06 (–0.15;0.26)	39
– SVM vs LR	0.03 (–0.20;0.26)	17
– ANN vs LR	–0.12 (–0.35;0.12)	27
– Other ML vs LR	–0.09 (–0.30;0.12)	46
High risk of bias		
– Any ML vs LR	0.34 (0.20;0.47)	137
– Tree vs LR	0.05 (–0.10;0.20)	26
– RF vs LR	0.41 (0.22;0.60)	20
– SVM vs LR	0.33 (0.19;0.48)	26
– ANN vs LR	0.71 (0.55;0.88)	25
– Other ML vs LR	0.31 (0.15;0.47)	40

Fig. 4. Differences in discriminative ability between LR and ML models, overall and according to risk of bias ($n = 282$ comparisons). When LR was compared with traditional statistical methods (discriminant analysis, Poisson regression, generalized estimating equations, generalized additive models), these methods were not included as "Other ML methods" and were thus excluded from this plot. LR, logistic regression; RF, random forest; SVM, support vector machine; ANN, artificial neural network.

- No evidence of superior performance of ***machine learning*** over ***logistic regression***.

Christodoulou. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction modelsJournal of Clinical Epidemiology. Volume 110, June 2019, Pages 12-22. 10.1016/j.jclinepi.2019.02.004,

Statistical model

- Survival analysis
  - Cox regression model provides a default framework for prediction of long-term prognostic outcomes.
  - Kaplan–Meier analysis provides a nonparametric method, but requires categorization of all predictors. It is the equivalent of cross-tables
  - Parametric survival models may be useful for predictive purposes because of their parsimony and robustness, for example at the end of follow-up

Statistical model

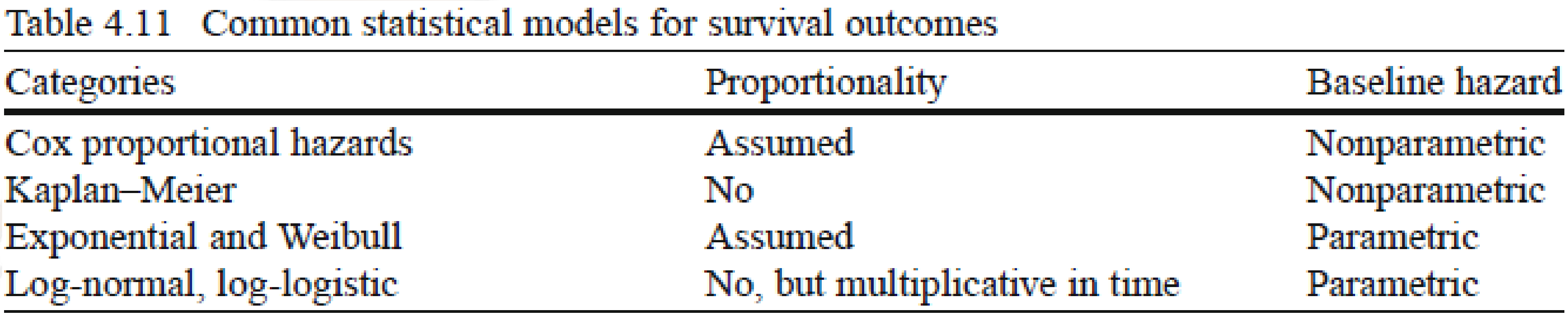


Table 4.11 Common statistical models for survival outcomes

Categories	Proportionality	Baseline hazard
Cox proportional hazards	Assumed	Nonparametric
Kaplan–Meier	No	Nonparametric
Exponential and Weibull	Assumed	Parametric
Log-normal, log-logistic	No, but multiplicative in time	Parametric

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Usual steps of CPM

- Modelling steps
  - Data inspection
    - Missing values
  - Coding of predictors
    - Continuous predictors; Combining categorical predictors
    - Restrictions on candidate predictors
  - Missing data
    - Simple imputation, multiple imputation (several methods)
  - Model specification
    - Appropriate selection of main effects?
    - Assessment of assumptions (distributional, linearity, and

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Usual steps of CPM

- Model estimation
  - Shrinkage included?
  - External information used?
  - Model performance appropriate statistical measures used?
  - Clinical usefulness considered?
- Model validation
  - Internal validation, including model specification and estimation?
  - External validation?

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Usual steps of CPM

- Validity
  - Internal: overfitting - sufficient attempts to limit and correct for overfitting?
  - External: generalizability - predictions valid for plausibly related populations?
- Model presentation
  - Format appropriate for audience?

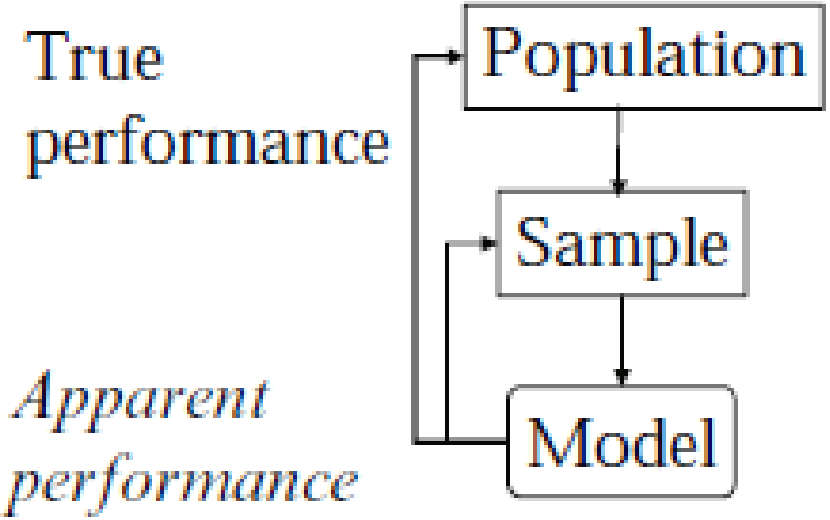
Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Model estimation

- Overfitting and Optimism
  - We are primarily interested in the validity of the predictions for new subjects, outside the sample under study
  - Overfitting causes optimism
  - **Overfitting** - the data under study are well described, but predictions are not valid for new subjects, usually accuracy is overestimated; a statistical model with too many degrees of freedom in the modelling process
  - **Optimism** – accuracy overestimation due to overfitting; true performance minus apparent performance
  - The solution is generally named “shrinkage” or penalization
  - Bootstrap resampling is a central technique to quantify optimism in internal model performance

Model estimation

- Overfitting and Optimism



Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Model estimation

- Overfitting and Optimism

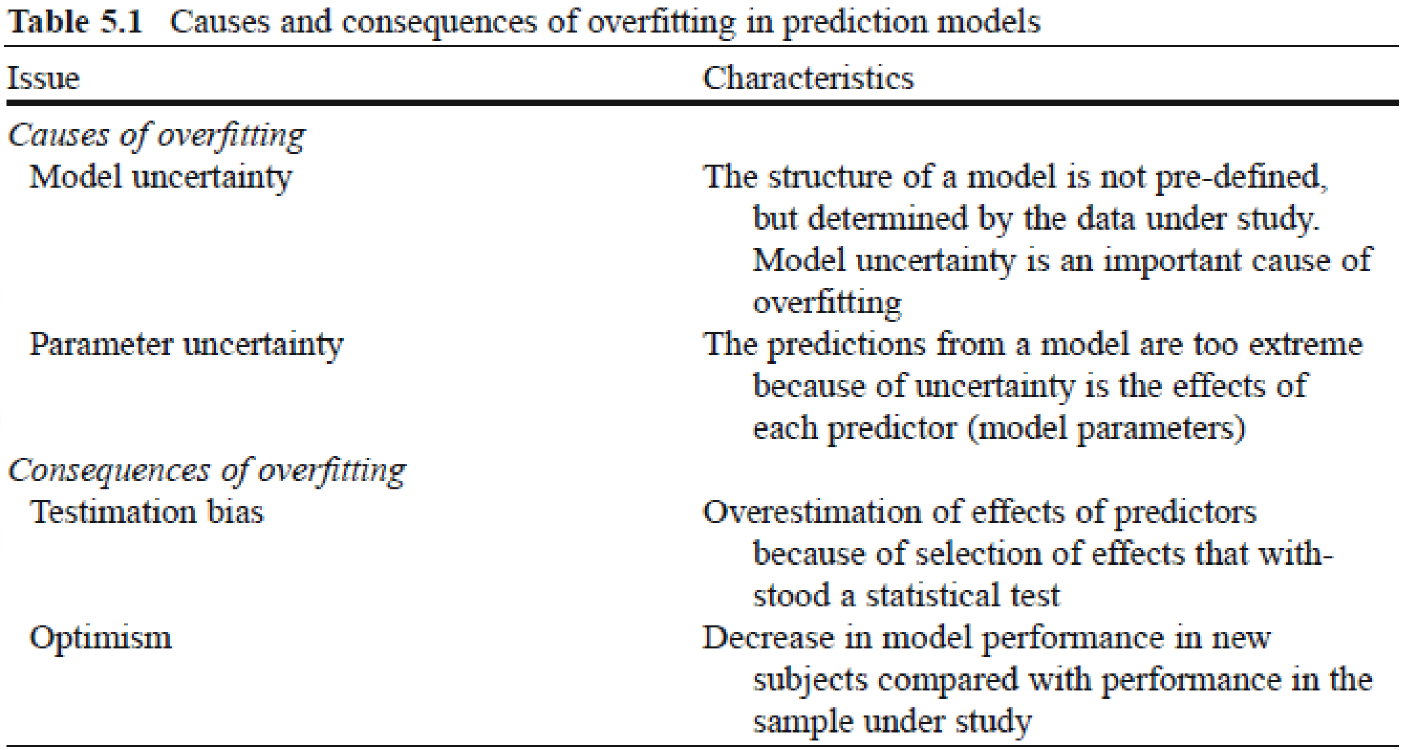


Table 5.1 Causes and consequences of overfitting in prediction models

Issue	Characteristics
Causes of overfitting	
Model uncertainty	The structure of a model is not pre-defined, but determined by the data under study. Model uncertainty is an important cause of overfitting
Parameter uncertainty	The predictions from a model are too extreme because of uncertainty is the effects of each predictor (model parameters)
Consequences of overfitting	
Testimation bias	Overestimation of effects of predictors because of selection of effects that withstood a statistical test
Optimism	Decrease in model performance in new subjects compared with performance in the sample under study

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Model estimation

- What is bootstrap?

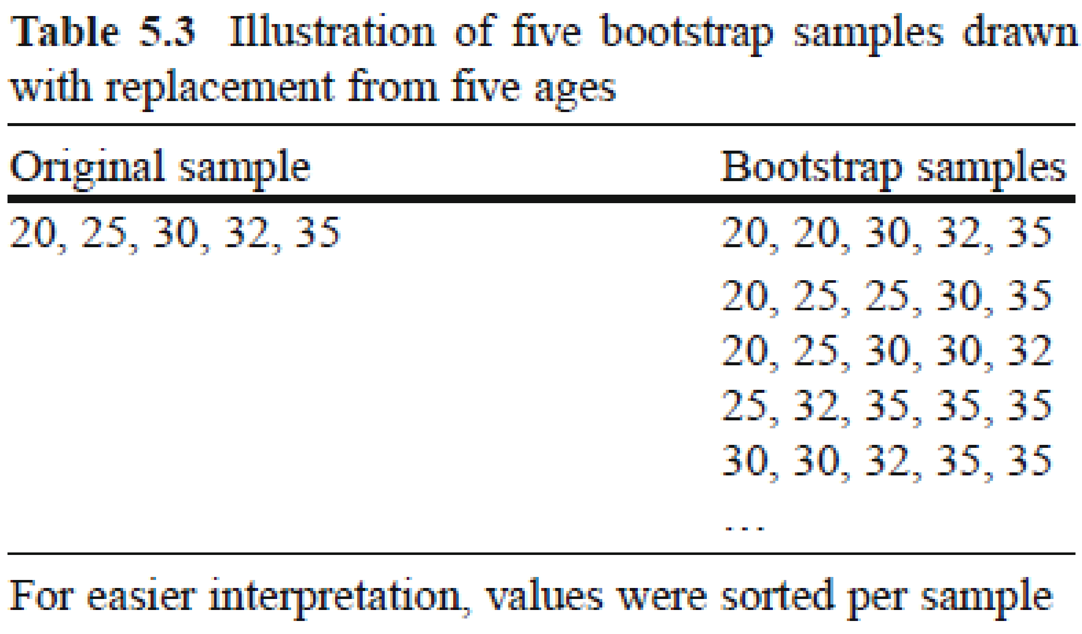


Table 5.3 Illustration of five bootstrap samples drawn with replacement from five ages

Original sample	Bootstrap samples
20, 25, 30, 32, 35	20, 20, 30, 32, 35
	20, 25, 25, 30, 35
	20, 25, 30, 30, 32
	25, 32, 35, 35, 35
	30, 30, 32, 35, 35
	…

For easier interpretation, values were sorted per sample

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Model estimation

- Bootstrap for calibration

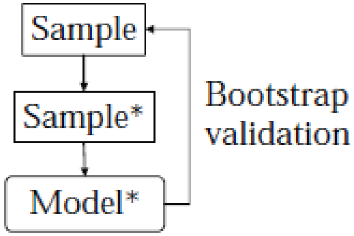


Fig. 5.7 Schematic representation of bootstrap validation for optimism correction of a prediction model. Sample* refers to the bootstrap sample that is drawn with replacement from the Sample (the original sample from an underlying population). Model* refers to the model constructed in Sample*

Optimism-corrected performance = Apparent performance in sample − Optimism

Optimism = Bootstrap performance − Test performance

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Model estimation

- Bootstrap for calibration

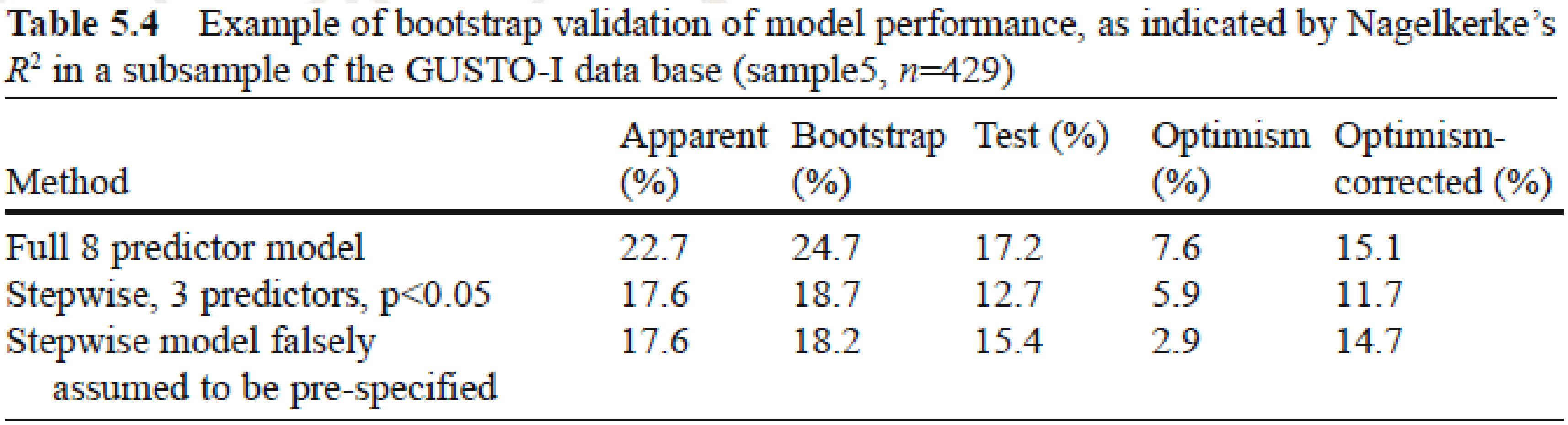


Table 5.4 Example of bootstrap validation of model performance, as indicated by Nagelkerke's R^2 in a subsample of the GUSTO-I data base (sample5, n=429)

Method	Apparent (%)	Bootstrap (%)	Test (%)	Optimism (%)	Optimism-corrected (%)
Full 8 predictor model	22.7	24.7	17.2	7.6	15.1
Stepwise, 3 predictors, p<0.05	17.6	18.7	12.7	5.9	11.7
Stepwise model falsely assumed to be pre-specified	17.6	18.2	15.4	2.9	14.7

Steyerbeg. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer in 2009.

Concluding

- There are recognized methodological standards that should be adhered to when developing and validating CPRs.
- The research design must follow the hypothesis question and each choice has it strong and weak points.
- Several analysis steps not usually included in other observational research must be considered, such as shrinkage, validation and calibration performance (apparent and corrected).

fim

Session 3
Steps in planning and conducting CPM research – Part 2

Pedro E A A do Brasil
pedro.brasil@fiocruz.br
2023