

Clinical Prediction Models

Ranstam J, Cook JA, Collins GS

Clinical prediction models (also known as prognostic models, risk scores) are mathematical equations that relates multiple predictors (risk factors, covariates) to the probability of having a disease or condition (diagnostic) or the probability that an event will happen in the future (prognostic)¹. In the field of surgery many models have been developed which predict outcome (e.g. mortality) following surgery. Well known prediction models include EuroSCORE II, P-POSSUM and the ACS NSQIP surgical risk calculator². Predictions from these models can be used for planning lifestyle changes, guiding therapeutic decisions, stratifying participants in randomized controlled trials or risk adjustment for comparing hospital or surgeon performance.

Whilst prediction models can be used to gain insight into causality of the outcome of interest, this is not the aim or indeed a requirement. Furthermore, not all predictors are causal but all causal factors can be a predictor. Very different sets of predictors can have a similar overall predictive performance.

There are clear steps to follow when introducing a new prediction model; model development, model validation and model impact^{3,4}. Model development involves selecting predictors and combining them into a multivariable model, typically using logistic or Cox regression. An important consideration at this stage is to carefully choose what predictors are examined, and limit the number of predictors relative to the sample size to avoid overfitting. Overfitting will reduce the predictive ability of the model when applied in new data. Model validation can be broadly separated into internal and external validation. Internal validation, which is part of the model development process, will quantify any optimism as a result of overfitting, using methods such as bootstrapping or cross-validation. External validation will evaluate model transportability on a different data (from which the model was developed). Both types of validation involve evaluating the performance of the model in terms of discrimination, calibration and sometimes clinical usefulness⁵. The final stage is model impact, by demonstrating whether using the model improves decision making⁶.

References

1. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann Intern Med* 2015;**162**(1): W1-W73.
2. Cook JA, Collins GS. The rise of big clinical databases. *Br J Surg* 2015; **102**:e93–e10.
3. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;**35**: 1925-1931.
4. Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating*. Springer: New York, 2009; XXVIII, 528.
5. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;**21**(1): 128-138.

6. Wallace E, Smith SM, Perera-Salazar R, Vaucher P, McCowan C, Collins G, Verbakel J, Lakhanpaul M, Fahey T. Framework for the impact analysis and implementation of Clinical Prediction Rules (CPRs). *BMC Med Inform Decis Mak* 2011;**11**: 62.

Word count 489