



Building models for prediction: are we good at it?

Marie Beauséjour¹ · Jean-Marc Mac-Thiong ²

Received: 29 September 2020 / Accepted: 30 September 2020 / Published online: 13 October 2020
© International Spinal Cord Society 2020

Models in general are representations of processes. When considering mathematical models, one may distinguish two main families of models: the one for which the various intervening factors remain stable over time, and dynamic models. The first family involves linear and non-linear mathematical representations commonly used as assumptions in statistical models. These models are those referred to in most studies whether experimental or observational. The second family is less often considered and more difficult to work with but is very often hypothesized in biological processes due to complexity.

Prediction models allow us to predict future behaviours or results that are as yet unseen or unmeasured [1]. They are mathematical representations of the evolution of a system from a wide variety of starting conditions, resulting in a set of numerical values called outcomes. Prediction models are important in various fields, including medicine, public health, rehabilitation, biology, and physics. A quick search at indexed journals in PubMed resulted in 18,395 papers with “prognostic model*” or “prediction model*” in the title or abstract over the past 20 years. Given this plethora of options, it can be challenging to hone in on the optimal model to guide clinical practice and decision. The 2020 editorial entitled: *Relationships, associations, risk factors and correlations: nebulous phrases without obvious clinical implications*, has made the point clear about the usefulness of studying relationships and associations “for one of two purposes: to predict an outcome or to identify causal mechanisms” [2]. The readers would find the 2019 editorial of Drs. Chamberlain and Brinkhof on causality enlightening

[3]. The present editorial will focus on predictive modelling for prognosis and especially on issues about validity and usefulness of such models.

Predictive modelling for prognosis uses patient characteristics to estimate the probability that a certain outcome will occur within a defined time period; models predict an outcome based on patterns from the past. This definition encompasses a clear consideration about timing of events for measuring predictors and outcomes. Prognostic modelling studies are intrinsically longitudinal, and models are built on data where the predictors precede the outcomes and where the outcomes are not present at enrolment [4]. Patients or participants that are included in the study should be at risk to develop the outcome of interest. Preferably, the sample is heterogeneous, including a wide range of values for the predictor variables. The best design for a prognostic modelling study is a prospective cohort [4, 5].

Developing a prediction model does not end with its construction. The building cycle comprises development, validation, eventually model updates, and even impact analysis [6–8]. The model development steps include internal validity evaluation to quantify any optimism in the predictive performance. This can be done through methods such as cross-validation (constructing several models using subsets of the input data available and evaluating them with the remaining subset of the data) and bootstrapping (creating random construction and evaluation subsets of the data from resampling with replacement). Useful metrics for this are calibration and discrimination. It is a necessary part of model development but insufficient to demonstrate the value of the model for use. After development, it is recommended to evaluate the performance of the model with participant data other than those used for the model development. It requires a new data set for which prediction is made, using the original model, and compared with the observed outcomes [6–8]. This new sample may be obtained with similar participants at a different time, in a different setting, or even with participants presenting with a different profile [7]. This constitutes external validation and

✉ Marie Beauséjour
marie.beausejour@usherbrooke.ca

¹ Department of Community Health Sciences, Faculty of Medicine and Health Sciences, University of Sherbrooke, Longueuil, QC, Canada

² Department of Surgery, Faculty of Medicine, University of Montreal, Montreal, QC, Canada

it concerns model generalizability. In case of poor performance, the model can be, under certain circumstances, updated or adjusted based on the validation data set [6].

Kent et al. reported from their experience that there are often conceptual misunderstandings about what can be inferred from a multivariable model from the perspectives of association, prediction, and causation [4]. When building and testing a model, we should ask ourselves what do we already know and what question are we trying to answer? Predictors/candidate predictors known? Functional form of model known? Coefficients known? [8]. In other words, what is already known about predictors/candidate predictors, on model specification and coding of predictors, or on model parameters estimation and added value of a new predictor. The prognostic modelling strategies will depend on the type of research question we are asking, and this will, in turn, drive the interpretation of findings. The framework developed by Kent et al. [4], by clarifying the conceptual positioning of the researcher, should help strengthen prognostic study design and interpretation.

Current approaches commonly used in predictive modelling are based on linear, additive, unidirectional reasoning. When testing the model against data, we are characterizing these premises as true. But is it really the functional form that we consider as the most valid representation to translate what we know about the evolution of the spinal cord system? Let us reconsider the second family of “mechanistic” models presented in the introduction. They use dynamic equations (often differential equations) because, at each step, their actual value for a parameter of interest depends on the past value of this same parameter. In other words, change depends on where you are in the process. They are complex models due to the non-linearity of the dynamics but they can also be complex because there are many elements involved, all interrelated with each other in a dynamic fashion [9]. When considering the mathematical models appropriate for predicting spinal cord system recovery, it is legitimate to ask the question if these complex models should be better explored.

Deficiencies in model development and validation ultimately lead to prediction models that are not or should not be used [7]. Prerequisites for clinical usefulness of a model include a clear rationale for developing or validating this model, coherence with current knowledge or triangulation, good predictive performance (which depends at least on the combination of discrimination and calibration), model applicability and capacity to guide important decisions. First, there is a clear advantage in testing candidate predictors that are measured early in the process and easy to measure, thus likely usable in the clinical setting. In addition, this reflection may involve considering the size and type of errors (systematic or not) or the error rate

in model predictions. In particular, one should look at the relative weight and consequences of false-positive or false-negative classification [8] and compare models according to potential patient and societal benefits (and harms). Predictions are most useful when decision making is complicated and when the clinical stakes are high. It is of great value to devote some time and attention to the presentation format of predictive modelling findings in order to translate them into useful information to guide practice and decision [8].

Finally, in order to be able to conduct a synthesis and critical appraisal of proliferating published prediction models in the field of spinal cord medicine and rehabilitation, the use of standards in reporting is helpful. Collins et al. have developed the TRIPOD statement, which is a guideline specifically designed for the reporting of studies developing and validating multivariable prediction models [7]. As this standard encourages reporting on the medical context and objectives of model elaboration, on clear definition and handling of outcomes and predictors, as well as on model specification and validation steps, we recommend its use for rigour and quality review of predictive modelling article submissions to *Spinal Cord*.

Acknowledgements The authors are grateful to Dr. Michel Bigras-Poulin, epidemiologist, retired professor at the Faculty of Veterinary Medicine of the Université de Montréal, for the enlightening discussions on this topic.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Dym C. Chapter 1 What is mathematical modeling? In: Principles of mathematical modeling. ProQuest Ebook Central. Elsevier Science; 2004. <http://ebookcentral.proquest.com/lib/kbdk/detail.action?docID=286730>. Accessed 24 Sep 2020.
2. Harvey L. Relationships, associations, risk factors and correlations: nebulous phrases without obvious clinical implications. *Spinal Cord*. 2020;58:1–2. <https://doi.org/10.1038/s41393-019-0396-8>.
3. Chamberlain JD, Brinkhof MWG. Using strong inference to answer causal questions in spinal cord injury research. *Spinal Cord*. 2019; 57:907–8. <https://doi.org/10.1038/s41393-019-0344-7>.
4. Kent P, Cancelliere C, Boyle E, Cassidy D, Kongsted A. a conceptual framework for prognostic research. *BMC Med Res Methodol*. 2020;20:172. <https://doi.org/10.1186/s12874-020-01050-7>.
5. van Kuijk SMJ, Dankers FJWM, Traverso A, Wee L. Preparing data for predictive modelling. In: Kubben P, et al., editors. Fundamentals of clinical data science; 2019. https://doi.org/10.1007/978-3-319-99713-1_6.
6. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation model updating, and impact assessment. *Heart*. 2012;98:691e8. <https://doi.org/10.1136/heartjnl-2011-301247>.
7. Collins GS, Reitsman JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for

- individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *J Clin Epidemiol*. 2015;68:112–21. <https://doi.org/10.1136/bmj.g7594>.
8. Steyerberg EW. Statistics for biology and health, clinical prediction models: practical approach to development, validation and updating. New York: Springer-Verlag; 2009. <https://doi.org/10.1007/978-0-387-77244-8>.
 9. Siegenfeld AF, Bar-Yam Y. An introduction to complex systems science and its applications. *Complexity*. 2020;6105872:1–16. <https://doi.org/10.1155/2020/610587>.