

# Clinical prediction model construction and evaluation: a challenge for clinical researchers

# Qiu-Yan Han, Zhi-De Hu

Department of Laboratory Medicine, The Affiliated Hospital of Inner Mongolia Medical University, Hohhot 010050, China *Correspondence to:* Zhi-De Hu. Department of Laboratory Medicine, The Affiliated Hospital of Inner Mongolia Medical University, Hohhot 010050, China. Email: hzdlj81@163.com.

Submitted Nov 23, 2019. Accepted for publication Nov 26, 2019. doi: 10.21037/atm.2019.11.145 View this article at: http://dx.doi.org/10.21037/atm.2019.11.145

In a recent issue of *Annals of Translational Medicine*, Zhou *et al.* comprehensively introduced the concept and applications of the clinical prediction model, as well as the methods and procedures of model construction (1). Some interesting and valuable statistical methods were introduced, such as nomogram, decision curve analysis (DCA), net reclassification index (NRI) and integrated discrimination index (IDI), and their R code was described in a step-bystep way. Undoubtedly, this special report provides a crucial guideline for researchers who are interested in the clinical model. The statistical aspects of the clinical prediction model have been clearly and comprehensively described in this special report. Here, I present some perspectives from clinical aspects.

# Variable selection

If the clinical model is a building, variables in the model are undoubtedly the bricks of the building. In the era of big data, variables can be obtained anywhere, such as routine laboratory tests, genetic information, history, signs, symptoms, complications, and microbiological findings. Some of the variable selection mothed from the statistical perspective has been introduced in the special report. Here, I added some principles from the clinical perspective.

# Economic consideration

Theoretically, models with more variables may have higher predictive accuracy. However, including more variables in a model results in a higher cost. Therefore, variables with low cost are preferred in clinical predicting model construction, such as history, signs, symptoms, routine hematological parameters, and liver and kidney function tests. If the total cost of the model is extremely high, especially it includes some non-conventional parameters (e.g., genetic information, lymphocyte subset analysis, circulating microRNAs), its clinical implications would be limited.

# Standardization

Some variables, especially laboratory tests, are well standardized. That means the results of the variable or test are comparable between different laboratories (e.g., hemoglobin, liver enzymes). These variables are preferred when building a model. However, for some tests or variables that are not well standardized, such as pleural interleukin-27 (IL-27) for tuberculosis pleurisy diagnosis, cautions should be raised, because the result of pleural IL-27 is greatly affected by the kit selected. In this case, adenosine deaminase (ADA) may be more preferred; even the diagnostic accuracy of ADA seems to be lower than IL-27 (2). For some symptoms or signs, their definitions are not clear, such as the New York Heart Association (NYHA) functional classification system in pulmonary arterial hypertension (PAH). The degree of agreement between clinicians is unsatisfied (3).

# Turn-around time (TAT)

For some emergency cases, such as chest or abdomen pain in the emergency department (ED). Rapidly differential diagnosis is crucial. When building a model with such a condition, TAT must be considered. Let's assume that

#### Page 2 of 3

a circulating microRNA measured by polymerase chain reaction (PCR) can provide diagnostic information for acute myocardial infarction (AMI) in chest pain patients who are admitted to ED, and thus including it into a predictive model is statistically reasonable. However, including this parameter into the model means: if we want to use the model to predict AMI in chest pain patients, results of circulating microRNA should be obtained firstly. The TAT of circulating microRNA is more than 3 hours. That means, to use this model, we need to wait for at least 3 hours. This model does not facilitate the early diagnosis of AMI, and its clinical efficiency is questionable. The TAT of the clinical predictive model depends on the variable(s) with the longest TAT.

#### Easily obtained

Some of the clinical variables, such as genetic information, are not easily obtained. Even for non-emergency cases, including such information in the model may impair the clinical utility of the model. We suggest the researcher used some easily obtained variables, such as symptoms, signs, history, and routine laboratory tests.

## Timepoint of variables

The model for predicting the prognosis of a given disease is of great value because the treatment or further long-term management approach is greatly affected by the prognosis of the patients. If the model aims to assist the clinicians in establishing a treatment approach, variables at baseline (before treatment) is preferred. It is a paradox to use a model that contains post-treatment variables (e.g., complications after surgery) as evident for treatment approach selection. If the models aim to guide further management (e.g., frequency of follow up after the operation, chemotherapy for postoperative lung cancer patients), post-treatment variables are acceptable.

The principals listed above are our perspective on variable selection in building the clinical model. It must acknowledge that, when building a model, there are tradeoffs between the predictive accuracy and cost, TAT, and standardization of the model. How to balance them is a challenge for model builders.

# Model evaluation: a long way from statistical significance to clinical efficiency

In the special report, the authors proposed some methods

for model evaluation with statistical methods, such as c-index, DCA, and calibration (1). These statistical methods are widely used in researches using the clinical model. The clinical practice is extremely complicated, rather than a model or algorithm consisted of some variables. Take heart failure (HF) diagnosis in acute dyspnea patients admitted to the ED as an example. It is well-known that some baseline characteristics can provide diagnostic information for HF, such as history, imaging details, routine laboratory tests. These characteristics can also be used to build a diagnostic model. Notably, if a scale is used to measure the diagnostic confidence of the clinicians who have known the baseline characteristics of patients, the area under the receiver operator characteristic (ROC) curve (AUC) is approximately 0.90 (4). Therefore, if the model has an inferior AUC than that of the scale, it is meaningless.

The improvement in diagnostic accuracy does not mean a better outcome (5). In other words, clinical utility is more attractive and valuable than predictive accuracy. The highlevel evidence to address the clinical utility of a predictive model is a randomized controlled trial (RCT). However, this type of study is rare. Some similar studies with single biomarkers have been performed, for example, the B-type natriuretic peptide (BNP) testing and clinical outcomes and health services in acutely dyspneic patients (6). In this RCT, 612 acute dyspnea patients admitted ED were randomized assigned to two groups with and without BNP testing. The outcomes, including admission rates, length of stay, mortality, and readmission rates, were compared to estimate the clinical utility of BNP testing. Using a similar strategy, we can use to evaluate the clinical utility of the diagnostic model develops. For the prognostic model, the clinical utility evaluation strategy is similar. For example, whether a model guided approach has superior clinical efficiency than the conventional treatment approach, just like the procalcitonin (PCT) guided guide initiation and duration of antibiotic treatment for patients with acute respiratory infections (7).

## Conclusions

The clinical model can facilitate the management of patients in clinical practice, and thus represents a hot clinical research strategy. The special report released by Zhou *et al.* comprehensively summarized the principals in clinical model construction and evaluation, thus bring us a comprehensive insight into the clinical predictive model. We believe that a perfect model should be easy to use, with high predictive accuracy, as well as acceptable clinical

#### Annals of Translational Medicine, Vol 8, No 4 February 2020

utility. Building such a model is a challenge for clinical researchers.

### **Acknowledgments**

None.

## Footnote

*Conflicts of Interest:* The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## References

 Zhou ZR, Wang WW, Li Y, et al. In-depth mining of clinical data: the construction of clinical prediction model with R. Ann Transl Med 2019;7:796.

**Cite this article as:** Han QY, Hu ZD. Clinical prediction model construction and evaluation: a challenge for clinical researchers. Ann Transl Med 2020;8(4):74. doi: 10.21037/ atm.2019.11.145

- 2. Porcel JM. Advances in the diagnosis of tuberculous pleuritis. Ann Transl Med 2016;4:282.
- Taichman DB, McGoon MD, Harhay MO, et al. Wide variation in clinicians' assessment of New York Heart Association/World Health Organization functional class in patients with pulmonary arterial hypertension. Mayo Clin Proc 2009;84:586-92.
- 4. Januzzi JL Jr, Camargo CA, Anwaruddin S, et al. The N-terminal pro-BNP investigation of dyspnea in the emergency department (PRIDE) study. Am J Cardiol 2005;95:948-54.
- Bossuyt PM, Reitsma JB, Linnet K, et al. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. Clin Chem 2012;58:1636-43.
- 6. Schneider HG, Lam L, Lokuge A, et al. B-type natriuretic peptide testing, clinical outcomes, and health services use in emergency department patients with dyspnea: a randomized trial. Ann Intern Med 2009;150:365-71.
- 7. Schuetz P, Briel M, Christ-Crain M, et al. Procalcitonin to guide initiation and duration of antibiotic treatment in acute respiratory infections: an individual patient data meta-analysis. Clin Infect Dis 2012;55:651-62.