



Moving Toward the Implementation of Precision Medicine Needs Highly Discriminatory, Validated, Inexpensive, and Easy-to-Use Prediction Models

Vincenzo Trischitta^{1,2} and
Massimiliano Copetti³

Diabetes Care 2020;43:701–703 | <https://doi.org/10.2337/dci19-0079>

Diabetes is one of the most challenging global health problems, affecting more than 400 million people (1). Although 10% of current global health care spending is devoted to diabetes (1), patients with diabetes remain at high risk of morbidity and mortality (2) mainly due to cardiovascular disease (3). This extremely heavy burden is likely to increase in the coming decades, especially considering the epidemic nature of the most common form, type 2 diabetes (4). It is, therefore, mandatory to address vigorously the negative impact of type 2 diabetes on vascular health and life expectancy. To this end, the availability of well-performing risk prediction models capable of identifying high-risk patients to be targeted with the most aggressive and most burdensome prevention strategies would play a pivotal role.

The study of Aminian et al. (5), published in this issue of *Diabetes Care*, presents several models (called Individualized Diabetes Complications [IDC] Risk Scores) able to estimate in obese patients with type 2 diabetes the risk of mortality and of long-term vascular complications, including coronary artery events, heart failure, and estimated glomerular

filtration rate (eGFR) <60 mL/min/1.73 m². Twenty-six baseline variables as potential predictors were modeled by time-to-event regression and random forest machine learning, an ensemble of survival regression trees grown on bootstrap resampling of the observations. The left-out data were then used to predict the error rate and, after permutation, to estimate the importance of a given predictor. In addition to the time-dependent area under the receiver operating characteristic curve (AUROC) and the calibration plot, a recently described index of prediction accuracy, which combines discrimination and calibration in a single value, was also used (6).

A total of 13,722 patients from the Cleveland Clinic database were analyzed retrospectively (i.e., 2,287 who underwent metabolic surgery and 11,435 propensity-matched nonsurgical individuals, with a 1:5 ratio). Discrimination ability as measured by the AUROC at 10 years in the surgical and nonsurgical groups, respectively, was 0.79 and 0.81 for all-cause mortality, 0.66 and 0.67 for coronary artery event, 0.73 and 0.75 for heart failure, and 0.73 and 0.76 for eGFR <60 mL/min/1.73 m². For readers

that are not acquainted with these numbers, AUROC (or the equivalent C-statistic) is the probability that the baseline risk predicted for an individual who will develop the event is greater than that predicted for a counterpart who will not develop the event; an AUROC of 0.79 means, therefore, that this probability equals 79%.

For an additional 12,816 Cleveland Clinic nonsurgical patients who had not been included in the training data set, a direct comparison between IDCs and the risk equations for complications of type 2 diabetes (RECODE) (7,8) was carried out. Data showed that, as compared with those from RECODE models, AUROCs from IDCs were modestly larger for all-cause mortality and heart failure (approximately 2% improvement) and clearly larger for low eGFR (17% improvement). Unfortunately, no comparison was performed for myocardial infarction.

Among the strengths of the study is the large sample size, comprising over 13,000 patients included in the analysis. Also noteworthy is the use of the random forest based on the bootstrap resampling and permutation strategy that simulate, de facto, natural variability and therefore

¹Department of Experimental Medicine, Sapienza Università di Roma, Rome, Italy

²Research Unit of Diabetes and Endocrine Diseases, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy

³Unit of Biostatistics, Fondazione IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy

Corresponding author: Vincenzo Trischitta, vincenzo.trischitta@uniroma1.it

V.T. and M.C. share equal authorship.

© 2020 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

provide an internal validation that somehow addresses the issue of false positive discoveries.

In addition, in line with the tradition of the Cleveland Clinic, which offers clinicians a vast library of risk calculators, all IDCs are readily available both on the web (<https://riskcalc.org>) and in a smartphone application (BariatricCalc). This study therefore falls within the framework of precision medicine and, in spite of the limits discussed below, represents a service for the entire diabetes community, making it possible to increase the level of medical decision personalization. Similarly, creditable and successful results have recently been achieved for predicting several chronic complications (7,8) and the risk of death (7–10) in patients with type 2 diabetes.

Some limitations regarding the study design may also be pointed out. As the authors clearly recognize, their study lacks an external, independent validation. That is to say, in their current version IDCs remain data driven, a major flaw when proposing prediction models as ready to be implemented in clinical settings. The authors should have attempted to validate their models at least in nonsurgical patients where validation would have been much simpler, as there are many established cohorts available worldwide for this purpose (7,8). The retrospective study design, which can cause several biases, is an additional major limitation. For example, the nonsurgical group could comprise individuals who for many reasons have not been assessed as suitable for metabolic surgery. This might call into question the usefulness of IDCs in conditions in which the right decision about the surgical therapeutic option needs to be made. Finally, with a median follow-up of only 3.9 years and with 75% of patients followed up for less than 6.1 years, the choice of 10 years as the time horizon for prediction appears bold.

We all dream of moving on to a context of fully established precision medicine to maximize effectiveness and to minimize both economic and personal costs of medical strategies. This dream will become reality when the medical management of each patient (or group of patients who share similar characteristics) is informed by well-functioning models capable of predicting individual risks and clinical trajectories. However, some essential needs must be met before a prediction model can be judged ready for clinical

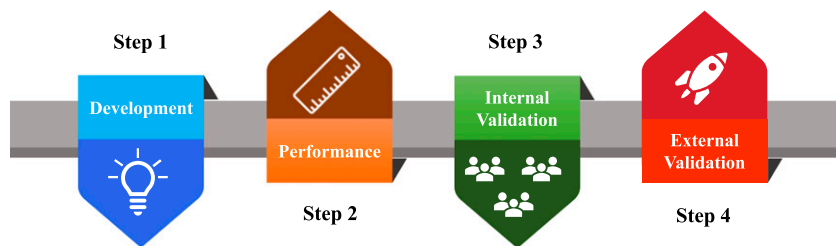


Figure 1—Development and validation of prediction models. Some methodological steps should be considered in developing and validating prediction models (12). With preference for a prospective study design, the model specification is followed by regression coefficients estimation (Step 1) using, ideally, shrinkage techniques, penalized estimation, or least absolute shrinkage and selection operator (LASSO) in order to limit overfitting (19). After that, a prediction model's performances (Step 2), including statistics for calibration and discrimination, should be measured avoiding overoptimism. Calibration-in-the-large, calibration slope, and C-statistic (or the equivalent AUROC) for discrimination are the most used indices to quantify a prediction model's quality (12,19,20). Step 3 refers to the validity of prediction within the population in which the model originated. This is known as reproducibility and is pursued through internal validation. Here, split sample validation is common but not efficient, while cross-validation and bootstrap resampling with replacement are to be preferred (12). Step 4 refers to generalizability or transportability of claims to populations that are different from those in which the model was created. To this end, the prediction model has to be externally validated in different clinical settings and/or in different genetic, environmental, and geographical backgrounds (19). Note that the greater the discrimination of the model, the greater its potential generalizability (18). Therefore, any attempt to improve discrimination, possibly up to outstanding AUROC or C-statistic values (i.e., >0.90), is welcome (18).

implementation. First, it should be based on prospective rather than retrospective cohorts in order to limit selection biases and reduce the risk of missing information. Second, well-performing models must be validated in additional, independent samples. This will ensure that prediction performance can be considered not only reproducible but also generalizable to different, though plausibly related, populations (11–13). Third, since approximately 80% of people with diabetes live in low- or middle-income countries (1) with limited health care resources, it would also be important that these tools be inexpensive, easy to use, and not time consuming. When all this is available, prediction models finally can be implemented in the real-life clinical setting where they were created. However, before it can be claimed that they can be used across different genetic, environmental, and geographical backgrounds, it must be shown that they are also transportable. This is not a feature to be taken for granted, as many studies have reported that in predicting cardiovascular disease and/or all-cause mortality, several validated prediction models, including the UK Prospective Diabetes Study (UKPDS) risk engine, the Framingham risk score, the Progetto Cuore, and the ENFORCE (Estimation of mORtality risk in type 2 diabetiC patiEnts) model, achieve worse performance if applied to patients from geographical regions

and/or clinical contexts different from those in which they were created (14–17). Notably, transportability of well-performing models may be ameliorated by improving their prediction accuracy, possibly up to outstanding discrimination, that is, AUROC or C-statistic values >0.90 (18).

In conclusion, creating useful prediction models is only possible if several methodological steps (Fig. 1) and social needs are taken into account. Once the model is established it can always be ameliorated by the addition of new clinical information and/or novel biomarkers derived, for example, from studies targeting genomic, metabolomic, and inflammation signatures. This will eventually improve the way we stratify disease-related risks, so to make possible moving toward the implementation of precision medicine.

Funding. This work was supported by the Italian Ministry of Health (Ricerca Corrente 2018–2020 to V.T.) and by the Italian Ministry of University and Research (PRIN 2015 to V.T.).

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

References

1. International Diabetes Federation. IDF Diabetes Atlas, 8th edition. Brussels, Belgium, International Diabetes Federation, 2017
2. World Health Organization. Global report on diabetes, 2016. Available from <https://www.who.int/diabetes/global-report/en/>. Accessed 6 April 2016

3. Rao Kondapally Seshasai S, Kaptoge S, Thompson A, et al.; Emerging Risk Factors Collaboration. Diabetes mellitus, fasting glucose, and risk of cause-specific death. *N Engl J Med* 2011;364:829–841
4. Ogurtsova K, da Rocha Fernandes JD, Huang Y, et al. IDF Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract* 2017;128:40–50
5. Aminian A, Zajichek A, Arterburn DE, et al. Predicting 10-year risk of end-organ complications of type 2 diabetes with and without metabolic surgery: a machine learning approach. *Diabetes Care* 2020;43:852–859
6. Kattan MW, Gerds TA. The index of prediction accuracy: an intuitive measure useful for evaluating risk prediction models. *Diagn Progn Res* 2018;2:7
7. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Yudkin JS. Development and validation of Risk Equations for Complications Of type 2 Diabetes (RECODE) using individual participant data from randomised trials. *Lancet Diabetes Endocrinol* 2017;5:788–798
8. Basu S, Sussman JB, Berkowitz SA, et al. Validation of Risk Equations for Complications of Type 2 Diabetes (RECODE) using individual participant data from diverse longitudinal cohorts in the U.S. *Diabetes Care* 2018;41:586–595
9. De Cosmo S, Copetti M, Lamacchia O, et al. Development and validation of a predicting model of all-cause mortality in patients with type 2 diabetes. *Diabetes Care* 2013;36:2830–2835
10. Copetti M, Shah H, Fontana A, et al. Estimation of mortality risk in type 2 diabetic patients (ENFORCE): an inexpensive and parsimonious prediction model. *J Clin Endocrinol Metab* 2019;104:4900–4908
11. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003;56:826–832
12. 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
13. Steyerberg EW, Harrell FE Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* 2016;69:245–247
14. van Dieren S, Peelen LM, Nöthlings U, et al. External validation of the UK Prospective Diabetes Study (UKPDS) risk engine in patients with type 2 diabetes. *Diabetologia* 2011;54:264–270
15. van Dieren S, Beulens JW, Kengne AP, et al. Prediction models for the risk of cardiovascular disease in patients with type 2 diabetes: a systematic review. *Heart* 2012;98:360–369
16. Guzder RN, Gatling W, Mullee MA, Mehta RL, Byrne CD. Prognostic value of the Framingham cardiovascular risk equation and the UKPDS risk engine for coronary heart disease in newly diagnosed type 2 diabetes: results from a United Kingdom study. *Diabet Med* 2005;22:554–562
17. Menzaghi C, Bacci S, Salvemini L, et al. Serum resistin, cardiovascular disease and all-cause mortality in patients with type 2 diabetes. *PLoS One* 2013;8:e64729
18. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York, John Wiley & Sons, 2000
19. Steyerberg E. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. 2nd ed. Basel, Switzerland, Springer, 2019
20. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21:128–138