The Royal New Zealand College of General Practitioners
Te Whare Tohu Rata o Aotearoa

VIEWPOINT

Precision health through prediction modelling: factors to consider before implementing a prediction model in clinical practice

Mohammad Z. I. Chowdhury MSc;[1] Tanvir C. Turin MBBS, PhD[1,2,3]

[1] Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Canada.
[2] Department of Family Medicine, Cumming School of Medicine, University of Calgary, G012F, Health Sciences Centre, 3330 Hospital Drive NW, Calgary, Alberta, Canada.
[3] Corresponding author. Email: chowdhut@ucalgary.ca

J PRIM HEALTH CARE
2020;12(1):3–9.
doi:10.1071/HC19087
Received 11 October 2019
Accepted 24 February 2020
Published 30 March 2020

ABSTRACT

INTRODUCTION: Precision medical practice emphasises early detection, improved surveillance and prevention through targeted intervention. Prediction models can help identify high-risk individuals to be targeted for healthy behavioural changes or medical treatment to prevent disease development and assist both health professionals and patients to make informed decisions. Concerns exist regarding the adequacy, accuracy, validity and reliability of prediction models.

AIM: The purpose of this study is to introduce readers to the basic concept of prediction modelling in precision health and recommend factors to consider before implementing a prediction model in clinical practice.

METHODS: Prediction models developed maintaining proper process and with quality prediction and validation can be used in clinical practice to improve patient care.

RESULTS: Aspects of prediction models that should be considered before implementation include: appropriateness of the model for the intended purpose; adequacy of the model; validation, face validity and clinical impact studies of the model; a parsimonious model with data easily measured in clinical settings; and easily accessible models with decision support for successful implementation.

DISCUSSION: Choosing clinical prediction models requires cautious consideration and several practical factors before implementing a model in clinical practice.

KEYWORDS: Prediction modelling; precision health; implement, clinical practice.

Introduction

Precision health care allows strategic use of data and time to study factors affecting an individual's unique health conditions to prevent, diagnose or treat disease. To prevent disease, the priority is to identify individuals at high risk of developing that disease. This is crucial for establishing individual intervention strategies. To implement effective prevention, health professionals need reliable tools to identify individuals free, but at risk of disease.

Modelling can help identify important risk factors contributing to outcomes; provide reasonable estimates about an illness's future course; and assist better care to avoid adverse events and prevent disease.[1] Prediction models are informative for both patients and clinicians as they provide quantifiable and readily interpretable predicted probabilities of an individual's risk for developing the disease.[2] For example, if the predicted risk of an individual is 0.20 for a particular disease, the probability of this

individual with specific risk factors developing the disease over some defined follow-up time interval is 20%. Various guidelines recommend incorporating prediction models in clinical practice due to their effectiveness.[3,4] However, not all models are developed robustly. Often flaws exist in prediction models.

In this article, we discuss what prediction modelling is, how it can help achieve precision health care and what to consider before deciding to implement a prediction model in clinical practice, using examples of prediction models to demonstrate how these specific factors are treated in the models. While the concept of prediction modelling is heavily statistical and may not be familiar to general audiences, we use plain language to help explain the topic. We also provide several references offering a thorough discussion of different aspects of prediction modelling for interested readers.

Prediction modelling

Prediction models are mathematical formulae or equations that combine multiple risk factors to express their relationship with an outcome and predict future outcomes. Prediction models vary depending on the nature of the relationship and types of variables. Building prediction models involves using datasets of individuals with known outcomes and applying the developed model to predict outcomes for future individuals. Regression methods, such as logistic regression (for binary outcomes) and Cox regression (for time-to-event outcomes), are frequently used to fit prediction models with clinical data. Numerous prediction models have been developed and used in clinical practice, public health, diagnostics, therapeutic decision-making and research. Several articles[1,2,5–7] and textbooks[8,9] discuss aspects of prediction modelling thoroughly and may offer interested readers a better understanding of this topic.

Purpose of prediction modelling

Clinical prediction models can help establish which risk factors determine the outcome and the degree of strength of association of each risk factor with the outcome, and predict the future of an outcome using specific values of certain risk factors. Predicting future outcomes has several applications, including detecting or screening high-risk individuals for asymptomatic disease to help prevent developing diseases through early intervention, and assisting in medical decision-making, which helps patients make informed choices regarding treatment.[10] Clinical prediction models also can assist with planning and quality management. Prediction models are informative for both patients and clinicians, providing a quantifiable and readily interpretable metric of an individual's risk for developing the disease. This information helps clinicians provide treatment recommendations and help patients make treatment decisions. The predictions obtained by these models should be absolute risk estimates of outcomes to guide forecasting for individuals. Patients are more concerned about their risk of getting an outcome in the future than their risk relative to other patients' risks. Relative risk estimates (odds ratios or hazard ratios) are used to get an absolute risk of the outcome.[1]

Precision health

Precision health is a priority research area. Precision medicine emphasises tailored prevention, diagnosis and treatment for individuals based on genetic, environmental and lifestyle factors.[11] Precision public health emphasises early detection, improved surveillance and targeted interventions.[11] Both focus on prevention through targeted intervention. Precision health is an emerging field that encourages disease prevention through earlier detection by monitoring health and disease based on individuals' risks.[12]

Prediction modelling and precision health

Disease prevention aims to end the costly cycle of disease management and its associated complications.[13] Primary prevention strategies are most effective when targeted to individuals at the highest risk. Identifying individuals at risk of developing a disease and identifying risk factors is crucial for developing individual intervention strategies. Screening people at the greatest risk of developing a disease maximizes the effectiveness of individualized preventive initiatives. Prediction models help identify individuals to be targeted. Targeted intervention has advantages over widespread generic intervention. In targeted intervention, people with

specific characteristics undergo screening as they (as a population) are at higher risk for a specific condition. For example, people with a family history of breast or ovarian cancer should be screened earlier and more intensively for these cancers. Targeted interventions provide the right intervention to the right person at the right time.

Widespread intervention involves a traditional priority group where all group members are suggested for screening (eg women aged 50–70 screened for breast cancer). Interventions achieve only average results when applied to everyone, and most people do not benefit. For example, if a model predicts a 20% risk that a person will develop hypertension, the observed frequency of hypertension should be 20 out of 100 people. If the intervention applies to everyone, most (80%) of the population do not benefit. Applying any intervention (medication, physical exercise, diet and lifestyle change) to everyone has the disadvantage that all are exposed to costs and harms (eg drug side-effects, anxiety). Prediction models should help direct targeted interventions so they focus on people who can benefit most and minimize unnecessary exposures to people unlikely to benefit.

There are concerns associated with prediction models. Incorrect prediction models may wrongly prioritise screened patients' access to high-risk health-care management programmes or new treatments. Prediction models may create racial and gender disparities, attributable either to people deciding how to build the models or to the data used in the model. There are many examples where prediction models produce discriminating results. A recent study showed how a health-care algorithm generated racial bias in predicting health risks.[14] The algorithm was used to predict patients with health conditions likely to lead to serious complications who would benefit most from additional assistance through high-risk care management. The flaw in the algorithm (which incorrectly turned some patients away from the care programme) led to a racial bias against treatment for ethnic minority patients. Careful scrutiny to avoid this type of algorithmic bias (which can be corrected) is needed before implementing prediction models in clinical settings.

Prediction models can also be used by public and advocacy groups. If readily available variables are used in models, the public can use the models to indicate their own risk for particular diseases, without consulting a doctor. Public use of prediction models allows screening for a whole population, which is not possible for clinicians alone. Public use of prediction models is complex and requires consideration of public awareness of the model's existence; understanding and interpreting the model's results; knowledge of the model's predicted outcomes; knowledge of the model's pros and cons; and knowledge of the implications of the model producing an incorrect decision.

Despite their advantages, applying prediction models in clinical practice is uncommon.[1,2] Reasons include the complexity of models intended for use in clinical settings; lack of sufficient validation and impact studies to make the models trustworthy; and inadequate understanding of the models and their predicted probabilities in decision-making by health professionals and patients.[1] Accurate, properly developed prediction models that are easy to use and have multiple validation and impact studies should be used in clinical settings, as they supplement other clinical information used in decision-making.

Example of a prediction model

The Framingham risk score[15] is a prediction tool widely used to predict individuals' 10-year cardiovascular disease (CVD) risk. This gender-specific risk score was developed based on age, total cholesterol, high-density lipoprotein (HDL) cholesterol, systolic blood pressure, antihypertensive medication use, diabetes and smoking to estimate the absolute CVD risk using data from the Framingham Heart Study. Individual CVD events (coronary heart disease, stroke, peripheral arterial disease and heart failure) were subsequently added to the Framingham risk score. A sample of size 8,491 mostly Caucasian people aged 30–74 years was used to develop the model, which showed good predictive performance (C-statistic, a measure to assess the predictive performance of the model, of 0.763 (men) and 0.793 (women)).

Considerations before implementing a model in clinical practice

Successfully implementing a prediction model often depends on how the model is supported and

recommended by experts in the field of application.[7] Generally, prediction models are implemented in clinical practice according to the recommendation of clinical practice guidelines. For example, the 2019 American College of Cardiology/American Heart Association guideline on the primary prevention of CVD recommended the race- and sex-specific Pooled Cohort Equation (atherosclerotic cardiovascular disease (ASCVD) Risk Estimator Plus) to estimate 10-year ASCVD risk and suggested asymptomatic adults aged 40–75 years should undergo this risk estimation before commencing pharmacological therapy to prevent the disease.[16] These guidelines were prepared by a group of scientists, researchers, clinicians and health professionals who are experts in the field. They thoroughly reviewed different aspects of the prediction models, discussed their pros and cons and assessed the available prediction models before making a recommendation. The end-users (clinicians) apply these prediction models in their clinical practice, according to the recommendation provided by the guidelines. As end-users, clinicians should have some understanding of these prediction models so they can communicate the models' results to their patients. They usually do not need to choose a particular prediction model to use in their practice, as this is provided by their regulatory bodies or guidelines.

We outline below five major aspects of a prediction model to consider before recommending its implementation in clinical practice.

1. Appropriateness of the model for the intended purpose

One priority of health and clinical research is to identify people who are at higher risk of developing an adverse health outcome, with the goal of targeting them for early preventive strategies and treatment. Numerous models have been developed to predict the future occurrence of disease.[17–19] Models can be developed to serve the same purpose using different patient characteristics. For example, a hypertension prediction model developed using the Chinese population may not serve the purpose of predicting hypertension in Caucasian patients. Models developed in secondary care may not be appropriate for primary care, as different settings have a different casemix.[2] Similarly, models developed using middle-aged and older adults may not be suitable for younger adults, as risk factors can be different for different age groups.[2] Models cannot be simply developed in one setting and applied to another. The participants used to develop and validate the model and the individuals for whom the model is recommended should be similar. For example, the Framingham risk score for men[15] can be used to estimate the general CVD risk of a 50-year-old Caucasian man in the US, but may not be suitable to estimate the same risk of a 50-year-old Asian man in China or the risk of a same-aged person in New Zealand. Before making recommendations, it is necessary to ensure the model is right for the intended purpose.

2. Adequacy of the model

Factors to be considered regarding the adequacy of the model include: (i) appropriateness of the variables included in the model; (ii) accuracy of the model; and (iii) sample size used to develop the model. The prediction model should contain all relevant variables chosen through clinical reasoning and by demonstrating statistical importance. For example, including body mass index (BMI) as a variable in predicting hypertension makes sense, while including 'tumour markers' (used in predicting cancer progression or recurrence) does not. A model without important variables related to the outcome fails to capture the true relationship that exists and provides inaccurate predictions. A model with too many variables creates complexity and raises the issue of generalisability.

A balanced model is preferred. Five to 20 variables often suffice to build an adequate prediction model.[8] The model also should predict the outcome accurately. A model without good predictive ability fails to identify patients who will and will not have the outcome in the future. A prediction model's accuracy is often assessed by the C-statistic; values for this measure range from 0.5 to 1.0, with higher values indicating better prediction. A C-statistic higher than 0.7 indicates a reasonable model; when it exceeds 0.8, models are considered strong.[20]

It is also important to know about the sample size used to develop the model. Models developed using small samples are often biased, less accurate and less generalisable than models developed with large

samples that can provide statistically valid and clinically useful results.[5] Determining adequate sample size is not straightforward. An event per variable (a ratio of the number of individuals with the outcome event to the number of candidate variables) of 10 is an often used and widely recommended sample size criterion for traditional prediction modelling.[21] The Framingham risk score[15] used seven variables, balancing simplicity and parsimony. Included variables were both clinically and statistically related to the outcome. The predictive accuracy of this model is good and it was derived using large sample size.

3. Validation, face validity and clinical impact studies of the model

The reliability and acceptability of prediction models largely depend on how well they perform in an external validation cohort, outside the derivation cohort used to develop the model. External validation uses data collected from a similar group of patients from a different setting to discover the accuracy and performance of a prediction model in a different patient population. Assessing a model's performance in a dataset not used to develop the model provides an unbiased estimate of the model's performance, as this external dataset is independent and can differ in many ways from the data used to develop the model (eg geographic location, time period, patient characteristics, investigators). Model performance in an external dataset is often worse than estimated from the development dataset due to real differences between the new setting and the development setting.[22,23] The generalisability of a model becomes stronger when the model is externally validated multiple times and demonstrates good predictive performance.[8] For example, the Framingham risk score[15] is widely used in many clinical settings due to its good predictive performance in multiple external validation studies. This increases the model's face validity and boosts confidence and trust among users.

Ideally, a prediction model will also have an impact study to assess whether the model improves clinical decision-making and patient health outcomes.[24] Impact studies also help identify factors (ease of use, acceptability) that can affect implementation in routine care.[2] Impact studies investigate whether the model improves clinical decision-making and patient health outcomes.[24] Conducting impact studies is expensive and requires significant effort, which is why only a few models have them.[24] Although not common, models with impact studies are assuredly effective in decision-making. Models with external validation and impact studies are recommended in clinical practice, as a model with accurate prediction is not beneficial if it is not generalisable or does not change behaviour.[2] Models must also fit the current context. A model developed years ago may not work for current patients for several reasons, including changes in risk factor distribution; availability of larger datasets on many risk factors; identifying new biomarkers; and developing new, improved prediction methodologies (eg machine learning tools). Temporal validation, where the model is validated in more recent individuals, and datasets can ensure a model's suitability for current use.

4. A parsimonious model with data that is easily measured in a clinical setting

Prediction models should be parsimonious (ie perform sufficiently well with the least amount of information) while explaining the data with good accuracy.[20] Models with too many variables are not practical, as information may not be available for all patients and may be costly to collect. Parsimonious models save the extra cost of collecting a large amount of information and time (eg clinician and patient time to collect information in routine visits) associated with measuring redundant variables. Simple parsimonious models are easy to interpret, generalise and use in practice.[8]

Prediction models should contain only variables that are clinically meaningful and easily measured in routine health-care visits to ensure their feasibility and applicability in clinical practice. For example, a hypertension prediction model should contain clinically important easily measured variables like BMI, smoking status, blood pressure, age, gender, marital status, ethnicity and parental history of hypertension but not total-to-HDL-cholesterol ratio, C-reactive protein (CRP) and other biomarkers, as they are not easily measured in a clinical setting. Biomarkers and other lifestyle-related factors can be useful in demystifying the complex cause of disease, but if they do not provide additional predictive information in determining

the future experience of a disease, it is better to exclude them from the prediction model. Often simple models with readily available variables perform better than complicated models with many variables.[25]

The inclusion of less easily measured variables, despite their significant contribution in prediction, will restrict a model's applicability in a clinical setting. For example, the Reynold Risk Score,[26] a more accurate CVD prediction model than the Framingham Risk Score,[27] is rarely used in clinical settings because it contains some less easily measured variables (eg high-sensitivity CRP) that are hard to obtain in routine primary care.

There is also little point in using even simple prediction models with only a few variables if they do not provide any additional information than the clinician's own risk assessment. Further, there are situations (eg a more serious outcome requiring in-depth investigation) that demand the use of complicated prediction models that have extraordinary predictive performance.

5. Accessible models with decision support for successful implementation

Prediction models should be easily accessible, easy to understand and usable during routine primary care consultations. Risk calculation should be simple and allow clinicians to generate the prediction. Using tools such as a risk chart or point-based risk score, derived from regression coefficients-based models, will be more meaningful in consultations. Mobile apps or web-based versions of these tools will make them more accessible to users (eg http://tools.acc.org/ASCVD-Risk-Estimator-Plus/#!/calculate/estimate/). Models can also be translated into a one-page checklist for patient self-assessment and can be made available in clinics.[10]

Prediction models can be presented with or without a decision recommendation. Models providing an explicit recommendation (eg reoccurrence of an event, treatment options, lifestyle modification) showed greater impact on clinical practice than models without a recommendation (providing predicted probability only), despite models without a recommendation allowing more room for clinicians to apply their clinical judgement.[24]

Supporting clinicians to use prediction models in their practice

For a prediction model to be useful in clinical practice, its end-users (clinicians and patients) must easily comprehend how the model works (which variables and risk factors were used to develop the prediction model, how the variables and the outcome were defined, how to interpret the results of the model). Knowing how a model works allows clinicians to make better intervention recommendations and patients to adhere to those recommendations more easily. Knowledge translation plays a vital role here. It would be very helpful for practicing clinicians to get training on the basics of how the models were derived and how they can be used as knowledge mobilization or health education. Information about a prediction model recommended by a proper guideline can be disseminated to its users through meetings, seminars, workshops and publications, and different health organisations and advocacy groups can help achieve this.

Conclusions

The high prevalence and global burden of disease make prevention and control strategies a top priority. Through prediction models, medical practitioners and researchers can better predict individual risks for a specific disease. This facilitates early detection and disease prevention and assists in making real-time decisions about the best way to care for patients and eventually helps achieve precision health care. While applying prediction models in clinical practice is uncommon, a well-developed model – one that considers the appropriateness, adequacy, parsimony and accessibility of the model before implementation – can help clinicians improve patient screening and ultimately prevent disease.

Competing interests

The authors declare no conflicts of interest.

Funding

This research received no grant support from any funding agency in the public, commercial or not-for-profit sectors.

References

1. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? BMJ. 2009;338:b375. doi:10.1136/bmj.b375
2. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606. doi:10.1136/bmj.b606
3. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III) final report. Circulation 2002;106(25):3143–421. doi:10.1161/circ.106.25.3143
4. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014;63:2889–934. doi:10.1016/j.jacc.2013.11.002
5. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. BMJ. 2009;338:b604. doi:10.1136/bmj.b604
6. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338:b605. doi:10.1136/bmj.b605
7. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381. doi:10.1371/journal.pmed.1001381
8. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer Science & Business Media; 2009.
9. Kuhn M, Johnson K. Applied Predictive Modeling. New York: Springer; 2013.
10. Lee YH, Bang H, Kim DJ. How to establish clinical prediction models. Endocrinol Metab (Seoul). 2016;31(1):38–44. doi:10.3803/EnM.2016.31.1.38
11. Prosperi M, Min JS, Bian J, Modave F. Big data hurdles in precision medicine and precision public health. BMC Med Inform Decis Mak. 2018;18(1):139. doi:10.1186/s12911-018-0719-2
12. Gambhir SS, Ge TJ, Vermesh O, Spitler R. Toward achieving precision health. Sci Transl Med. 2018;10(430):eaao3612. doi:10.1126/scitranslmed.aao3612
13. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/ PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2018;71:e127.
14. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447–53. doi:10.1126/science.aax2342
15. D'Agostino RB, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care. Circulation. 2008;117(6):743–53. doi:10.1161/CIRCULATIONAHA.107.699579
16. Arnett DK, Blumenthal RS, Albert MA, et al. ACC/AHA guideline on the primary prevention of cardiovascular disease. J Am Coll Cardiol. 2019;2019:26029.
17. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst. 1989;81(24):1879–86. doi:10.1093/jnci/81.24.1879
18. Stevens RJ, Kothari V, Adler AI, et al. The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). Clin Sci (Lond). 2001;101(6):671–9. doi:10.1042/cs1010671
19. Lindström J, Tuomilehto J. The diabetes risk score. Diabetes Care. 2003;26(3):725–31. doi:10.2337/diacare.26.3.725
20. Hosmer DW, Jr, Lemeshow S, Sturdivant RX. Applied Logistic Regression. New York: John Wiley & Sons; 2013.
21. Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. J Clin Epidemiol. 2016;76:175–82. doi:10.1016/j.jclinepi.2016.02.031
22. Steyerberg EW, Bleeker SE, Moll HA, et al. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. J Clin Epidemiol. 2003;56(5):441–7. doi:10.1016/S0895-4356(03)00047-7
23. Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol. 2003;56(9):826–32. doi:10.1016/S0895-4356(03)00207-5
24. Kappen TH, van Klei WA, van Wolfswinkel L, et al. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn Progn Res. 2018;2(1):11. doi:10.1186/s41512-018-0033-6
25. Paynter NP, Cook NR, Everett BM, et al. Prediction of incident hypertension risk in women with currently normal blood pressure. Am J Med. 2009;122(5):464–71. doi:10.1016/j.amjmed.2008.10.034
26. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA. 2007;297(6):611–9. doi:10.1001/jama.297.6.611
27. Cook NR, Paynter NP, Eaton CB, et al. Comparison of the Framingham and Reynolds Risk scores for global cardiovascular risk prediction in the multiethnic Women's Health Initiative. Circulation. 2012;125:1748–56. doi:10.1161/CIRCULATIONAHA.111.075929