Biostatistics behind risk prediction models

A risk prediction model is a mathematical equation that uses patient risk factor data to estimate the probability of a patient experiencing a healthcare outcome. Risk prediction models are used throughout medical practice for a variety of purposes such as predicting the development of a disease, predicting response to treatment, or predicting patient prognosis. Risk prediction modeling is at the forefront of improving the quality of care, reducing costs, and improving population health overall.

The authors intend to share a few observations that they had while going through the risk prediction models research articles:[1]

1. Importance of including likelihood ratios (LRs)

In studies such as these, it is important to include likelihood ratios (LRs) in addition to sensitivity, specificity, and predictive values. LRs expressed in an easy-to-comprehend language are more useful to clinicians in providing a better interpretation and adoption into operational practice. Eventually, the value of a score/test to influence clinical management decisions will depend upon its ability to alter the pre-test probability of a target condition into, what we call, the post-test probability. A positive LR >10 or a negative LR <0.1 are considered to exert highly significant changes in probability, which is, in turn, sufficient to alter clinical management.

2. The c statistic or receiver operating characteristic curves (ROC) may not be optimal in assessing models that predict future risk or stratify individuals into risk categories

ROC curves compare sensitivity versus specificity across a range of values for the ability to predict a dichotomous outcome. The area under the ROC curve (AUROC) is another measure of test performance. However, all these parameters are not intrinsic to the test and are determined by the clinical context in which the test is employed. The *c* statistic or AUROC uses the test characteristics of sensitivity and specificity to differentiate diseased from healthy patients and is a popular diagnostic test tool. However, it may not be optimal in assessing models that predict future risk or stratify individuals into risk categories. When the goal of a predictive model is to categorize individuals into risk strata, the assessment of such models should be based on how well they achieve this aim. To compare global model fit, use a measure based on the log-likelihood function, such as the Bayes information criterion, in which lower values indicate better fit and a penalty is paid if the number of variables is increased. Cook *et al*.[2,3] gave four suggestions for comparison of models for risk prediction, which are very apt and valuable for readers working on risk prediction models.

3. **Assessing the value of risk predictions using risk stratification tables**

A novel approach to assessing the value of adding a new marker to a risk prediction model is called the risk stratification approach.[2-4] This involves cross-tabulating risk predictions on the basis of models with and without the new marker, and has been widely adopted in the literature. It is suggested that the readers look into this as well.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

Kanica Kaushal[1], Sunil K. Raina[2]

[1]Department of Clinical Research and Epidemiology, Institute of Liver and Biliary Sciences, New Delhi, India, [2]Community Medicine, Dr. R. P. Government Medical College, Tanda, Himachal Pradesh, India

Address for correspondence: Dr. Sunil K. Raina, Professor and Head, Community Medicine, Dr. R. P. Government Medical College, Tanda, Himachal Pradesh, India.
E-mail: ojasrainasunil@yahoo.co.in

References

1. Karna ST, Gouroumourty R, Ahmad Z, Trivedi S, Thaware P, Singh P. Performance of prognostic scores in prediction of 30-day postoperative mortality in COVID-19 patients after emergency surgery: A retrospective cohort study. J Postgrad Med 2022;68:199-206.
2. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928-35.
3. Cook NR, Buring JE, Ridker PM. The effect of including C-reactive protein in cardiovascular risk prediction models for women. Ann Intern Med 2006;145:21-9.

4. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. Ann Intern Med 2008;149:751-60.

Received: 11-12-2022 **Accepted:** 22-01-2023
Published: 31-05-2023

Access this article online

Quick Response Code:

Website: www.jfmpc.com

DOI: 10.4103/jfmpc.jfmpc_2406_22

How to cite this article: Kaushal K, Raina SK. Biostatistics behind risk prediction models. J Family Med Prim Care 2023;12:1016-7.