Journal of Clinical Epidemiology 69 (2016) 89–95

Journal of Clinical Epidemiology

Prediction impact curve is a new measure integrating intervention effects in the evaluation of risk models

William Campbell[a,1], Andrea Ganna[b,c,1], Erik Ingelsson[c], A. Cecile J.W. Janssens[a,d,*]

[a]Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA, USA
[b]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, PO Box 281, SE-171 77 Stockholm, Sweden
[c]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, BOX 1115, 751 41 Uppsala, Sweden
[d]Department of Clinical Genetics, Section Community Genetics, EMGO Institute for Health and Care Research, VU University Medical Center, PO Box 7057, 1007 MB Amsterdam, The Netherlands

Accepted 16 June 2015; Published online 25 June 2015

Abstract

Objective: We propose a new measure of assessing the performance of risk models, the area under the prediction impact curve (auPIC), which quantifies the performance of risk models in terms of their average health impact in the population.

Study Design and Setting: Using simulated data, we explain how the prediction impact curve (PIC) estimates the percentage of events prevented when a risk model is used to assign high-risk individuals to an intervention. We apply the PIC to the Atherosclerosis Risk in Communities (ARIC) Study to illustrate its application toward prevention of coronary heart disease.

Results: We estimated that if the ARIC cohort received statins at baseline, 5% of events would be prevented when the risk model was evaluated at a cutoff threshold of 20% predicted risk compared to 1% when individuals were assigned to the intervention without the use of a model. By calculating the auPIC, we estimated that an average of 15% of events would be prevented when considering performance across the entire interval.

Conclusion: We conclude that the PIC is a clinically meaningful measure for quantifying the expected health impact of risk models that supplements existing measures of model performance.

Keywords: Prediction impact curve; AUC; Risk model; Predictive model; Coronary heart disease; Predictive ability

1. Introduction

The performance of risk models is evaluated in terms of clinical validity and clinical utility. To assess the clinical validity and utility of risk models, several traditional and novel measurements are available. (See [1] for a review.) The most frequently used measure of clinical validity is the area under the receiver operating characteristic (ROC) curve (AUC or auROC) [2]. AUC quantifies the ability of a risk model to discriminate between individuals who will or will not manifest the outcome of interest (referred to as events and nonevents in this article). A risk model with a higher AUC will be able to better separate the predicted risk distribution curves of events and nonevents. AUC lacks an important parameter: the incidence of the outcome in the population, and thus, it is not well suited for clinical decision making [3].

Measures of clinical utility indicate whether the risk model has a health or health care impact. One aspect of clinical utility is whether adding risk factors to a risk model changes medical decisions by reclassifying individuals into different risk categories. Measures of reclassification quantify this utility of updating a risk model [4]. The most commonly reported measure of reclassification, net reclassification improvement (NRI), considers reclassification separately for events and nonevents. NRI is the sum of the net reclassification in each of the groups because of updating the model. NRI has been criticized for systematically inflating the apparent performance of a new predictor [3,5].

Funding: This work was supported by The European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE Consortium, grant agreement HEALTH-F4-2007-201413, the Swedish Research Council (project grant no. 2012-1397), the Swedish Heart-Lung Foundation (project grant no. 20120197), the National Cancer Institute at the National Institutes of Health (grant number HHSN261201200425P), and the European Starting grant of the European Research Council (#310884).

Conflicts of interest: None.

[1] Shared first author.

* Corresponding author. Tel.: +1 404 727 6307; fax: +1 404 727 8737.
E-mail address: cecile.janssens@emory.edu (A.C.J.W. Janssens).

http://dx.doi.org/10.1016/j.jclinepi.2015.06.011
0895-4356/

What is new?

- The prediction impact curve is a new metric that measures the performance of risk models in a clinically meaningful way.
- The prediction impact curve is used to measure the percentage of events that are preventable when using a risk model to assign a preventive or therapeutic intervention.
- When considering the utility of risk models in clinical practice, researchers should focus on developing decision-analytic measures of performance.

AUC and NRI are widely reported summary statistics, but they both have notable shortcomings and lack clinically meaningful interpretations [6–10]. AUC is the probability that a randomly selected event has a higher predicted risk than a randomly selected nonevent, and NRI is the sum of the net percentage of events that move to a higher risk category and the net percentage of nonevents that move to a lower category.

Recent emphasis has been placed on the development of decision-analytic measures of predictive ability when considering the use of risk models in clinical practice [3]. Accordingly, we propose a new plot and measure, the prediction impact curve (PIC), and the area under the PIC (auPIC), which are intended as clinically meaningful measures of the expected population health impact that results from using a risk model to assign a preventive or therapeutic intervention to high-risk individuals.

In this article, we explain how the PIC is constructed, demonstrate its interpretation, and propose to calculate the area underneath as a summary statistic. Second, we investigate the properties of the auPIC in relation to the factors that determine the area. Finally, we apply the PIC to data from the Atherosclerosis Risk in Communities (ARIC) Study and illustrate how this graphical approach can be used to estimate the percentage of coronary heart disease (CHD) events prevented when using a risk model to assign statin treatment.

2. Materials and methods

2.1. Terminology and assumptions

The PIC can be constructed for any risk model and any outcome, such as disease, prognostic outcomes, or treatment response, but for reasons of readability, this article considers the prediction of future disease in a cohort of individuals that is followed over time. Individuals who develop the disease during follow-up are referred to as events, others are nonevents. We assume the availability of an intervention that has a certain preventive effect, indicated by the preventive fraction (PF), and which is given to individuals with the highest predicted risk. Throughout this article, we will refer to "high-risk" individuals, but it is important to note that our definition of high risk is determined by predicted risk scores and is allowed to vary by changing the risk threshold at which the high-risk group is defined.

The PF is assumed to be independent of the individual's predicted risk. This assumption was necessary because of the near impossibility of determining the covariance between PF and predicted risk because a given risk score could be the result of any one of a large number of possible covariate patterns in the risk model.

2.2. Constructing the prediction impact curve

The PIC plots the percentage of the population defined as high risk (x-axis) against events prevented (y-axis), which is the percentage of events seen in an untreated population that is prevented when the intervention is given to the high-risk group (Fig. 1).

The percentages of events prevented are obtained for every possible risk threshold that could be used to define the high-risk group. The size of the high-risk group is the percentage of all individuals with predicted risks that are higher than the risk threshold. The smallest increment in risk threshold and high-risk group size is achieved by adding the next ranked individual to the high-risk group, when individuals are ranked by decreasing predicted risks.

Three parameters are necessary to calculate the PIC: the sensitivity of the risk model at the risk threshold, the PF of the intervention, and the cumulative incidence of the disease in the population over a specified period of time.

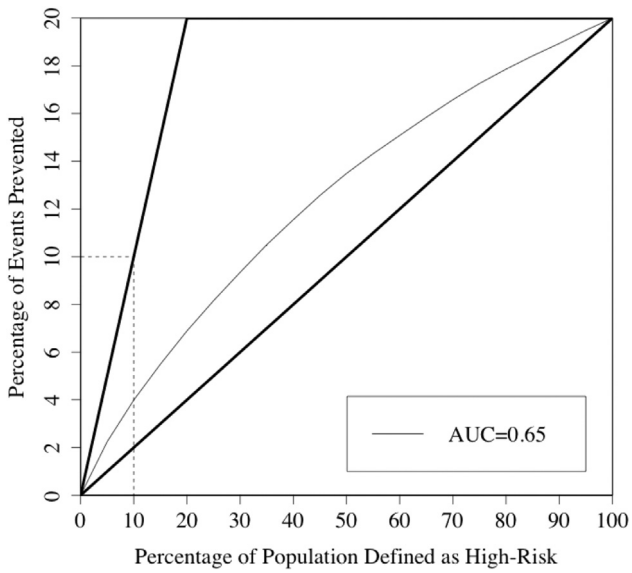

Fig. 1. Prediction impact curve for one risk model. The plot presents a scenario in which disease incidence was 20% and the intervention had a PF of 20%. The bold lines represent the theoretical maximum and minimum prediction impact curves. AUC, area under the receiver operating characteristic curve; PF, preventive fraction.