Implementing Cardiovascular Risk Prediction in Clinical Practice: The Future Is Now

Kunal N. Karmali, MD, MS; Donald M. Lloyd-Jones, MD, ScM

Risk prediction equations have been a cornerstone of cardiovascular disease prevention strategies for 2 decades. These equations serve as tools to convert data on multiple risk factors into a summary estimate of a person’s likelihood of experiencing a cardiovascular event over a given period. The first widely used cardiovascular risk prediction equation was the Framingham Risk Score (FRS), developed from the country’s first longitudinal cardiovascular cohort study. Eventual adoption of the FRS into the Third Report of the National Cholesterol Education Program’s Adult Treatment Panel (ATP-III) cholesterol guidelines in 2001 firmly established absolute risk assessment as an integral part of primary prevention, operationalizing the widely accepted paradigm that more intensive prevention efforts, specifically drug therapy, should be directed to those at higher risk.¹

In 2013, the American College of Cardiology and American Heart Association released updated clinical practice guidelines for the treatment of blood cholesterol to reduce atherosclerotic cardiovascular disease event risk.² These guidelines reaffirmed the risk-based prevention paradigm but moved one step further by eliminating cholesterol goals, instead identifying evidence-based risk thresholds to define groups with net benefit from statin therapy, in order to guide clinician–patient decision making in primary prevention. The prominent role of absolute risk assessment in these guidelines led to an intense focus on the new cardiovascular risk prediction equations recommended by these guidelines, the Pooled Cohort Equations (PCE).³

Since release of the 2013 American College of Cardiology/American Heart Association prevention guidelines, there have been numerous studies evaluating the performance of the PCE in different settings; reported results have been mixed, and findings have been heavily influenced by diverse and contentious methodological approaches in those reports.⁴–⁶ Some analyses have identified overprediction of risk with the PCE,⁷–¹¹ while others have found acceptable calibration, particularly at clinically relevant risk levels near decision thresholds.¹²–¹⁵ The prevailing uncertainties have led to calls for transformative changes in the way risk prediction algorithms are developed and validated.⁶,¹⁶ One potential approach is to move away from population-based cohort studies toward contemporary and real-world populations from electronic health records (EHRs) that reflect current trends in racial diversity, risk factor prevalence, preventive medication use, and disease incidence.

Yet, the use of EHRs as a tool for clinical research is still in its infancy, and few health systems have follow-up long evaluations to advance the field.¹⁷

In this issue of JAHA, Wolfson et al report a new analysis that evaluates the performance of 2 cardiovascular risk prediction equations in an integrated healthcare system with a mature and comprehensive EHR.¹⁸ The investigators analyzed data from 84,116 adults aged 40 to 79 years who were part of the HealthPartners system in Minnesota from 2001 to 2011 to determine the discrimination and calibration of the 2007 general FRS equations¹⁹ and the PCE.³ Using accepted methods for recalibration, the investigators also evaluated the performance of refitted FRS and PCE models within their system. In keeping with the pragmatic nature of EHR-based studies, the authors used risk factor measurements that were collected (or imputed values for those not collected) as part of routine clinical care and identified cardiovascular events by insurance claims data and state vital records that are included in the HealthPartners system.

Importantly, the authors found that, in this real-world EHR cohort, both the published and refitted FRS and PCE produced relatively accurate risk predictions. Specifically, the original FRS had a C-index of 0.740 (95% CI, 0.724–0.755) and a calibration statistic of 9.1 (P=0.028), while the PCE had a

The opinions expressed in this article are not necessarily those of the editors or of the American Heart Association.

From the Division of Cardiology, Departments of Medicine (K.N.K., D.M.L.-J.) and Preventive Medicine (D.M.L.-J.), Northwestern University Feinberg School of Medicine, Chicago, IL.

Correspondence to: Donald M. Lloyd-Jones, MD, ScM, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 N Lake Shore Dr, Suite 1400, Chicago, IL 60611. E-mail: dlj@northwestern.edu

Am Heart Assoc. 2017;6:e02254. DOI: 10.1161/JAHA.117.006019.

© 2017 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
C-index of 0.747 (95% CI, 0.727–0.768) and a calibration statistic of 43.7 (P<0.001). Furthermore, visual assessment of both calibration plots was acceptable. Not surprisingly, calibration was better with refitted models but results were qualitatively similar.

A key strength of this analysis is the inclusive selection criteria used by the investigators that produced a real-world and representative primary care population. Overly restrictive selection criteria in such validation studies can lead to bias, an underappreciated threat that has plagued previous attempts to study these equations in EHR cohorts. Additionally, the authors employed robust, multiple imputation methods to account for missing lipid data, a reality of working with real-world EHR data.

There are, however, some limitations worth noting. First, the studied population was quite similar in racial and demographic characteristics to the cohorts from which the FRS and PCE were derived. This likely explains the minimal demographic characteristics to the cohorts from which the studied population was quite similar in racial and working with real-world EHR data.

Second, because of the authors’ reliance on administrative for outcome assessment, the risk for misclassification exists, particularly for the outcomes of peripheral arterial disease and heart failure predicted by the general FRS equations. While this decision resulted from pragmatic and defensible considerations, future research will be needed to fully appreciate how this compares with the standardized methods used for outcome adjudication in many population-based cohorts. Third, although the general FRS equations have been available for nearly a decade, they have not been incorporated into any clinical practice guideline, and they contain heterogeneous atherosclerotic and nonatherosclerotic clinical outcomes. Therefore, risk estimates from this FRS do not align with any specific guideline recommendation.

In the 2013 American College of Cardiology/American Heart Association cholesterol guidelines, for example, the 10-year absolute risk threshold of 7.5% was specifically identified to mark a risk level where clinical trial data demonstrated that benefits of statin treatment for fatal and nonfatal atherosclerotic events clearly outweighed known risks of adverse events.

Finally, the analyses by Wolfson et al focus only on statistical metrics that evaluate model performance (discrimination and calibration) but do not indicate the downstream implications of these estimates on treatment decision-making. From a clinical perspective, calibration in particular is a visual exercise more than a statistical exercise. P-values for calibration are notoriously sensitive to sample size, and they do not indicate in which part of the risk spectrum any miscalibration may be occurring. Obviously, good calibration is most important near potential decision thresholds, and it is less important (or even irrelevant) at the extremes of the risk distribution. In the Wolfson et al analysis, for example, the PCEs were very well calibrated at low and moderate risk ranges, and overpredicted only in ranges above the clinical decision threshold of 7.5%, where “overprediction” is far less important, and may actually be a function of the application of risk-reducing therapies during follow-up that altered the predicted natural history of atherosclerotic cardiovascular disease risk. At lower levels of risk, such as 10-year risk levels of 5% to 7.5% and 7.5% to 10%, predicted event rates for the FRS were lower than the observed rates (6.1% predicted versus 6.5% observed and 8.6% predicted versus 10.4% observed, respectively). In contrast, the PCE slightly overpredicted risk at these same thresholds (6.1% predicted versus 5.6% observed and 8.6% versus 7.4% observed, respectively).

While the former might have better calibration, one might accept a more sensitive risk estimator from a public health perspective, particularly when considering the use of safe, effective, and low-cost medications such as statins. These limitations notwithstanding, the analysis by Wolfson et al is an important and valuable demonstration of the successful application of the FRS and PCE to a modern EHR system and should hopefully address uncertainties about the relevance of these equations in the contemporary era.

As clinical practice guidelines continue to move toward personalized treatment recommendations that are tailored to the unique benefit–harm assessments of a given patient, integration of clinical risk prediction equations will remain essential for guiding absolute risk assessment. Continued progress in health information exchanges and the establishment of standards for data harmonization, data quality, and electronic outcome assessment may one day lead to a nationwide electronic cohort capable of supporting ongoing refinement of risk prediction equations using real-world clinical data. However, until that time, we will likely save far more lives and prevent many more events by focusing on implementation of existing guideline-linked equations such as the PCE, with decision-support algorithms, in EHR platforms.

Predicting the future is an inherently imperfect science, but we must not forget that quantitative risk assessment is just the start, not the end, of a treatment decision. Risk estimates must be contextualized by clinicians for patients during a shared treatment discussion. Although recent years have seen great interest in the accuracy of cardiovascular risk prediction equations, there remains uncertainty over whether use of any cardiovascular risk estimate in clinical practice actually improves cardiovascular outcomes, and there are very limited data on how to best present this information for clinical decision making. Ultimately, analyses such as those
by Wolfson et al should serve to remind us that currently available risk prediction equations, even those derived from “historical” cohorts, remain applicable today. Now, we must continue the difficult work of identifying the best strategies for implementing these tools in practice to end the epidemic of cardiovascular disease in the population.

Disclosures

None.

References


Key Words: Editorials • health services research • primary prevention • risk assessment • risk prediction
Implementing Cardiovascular Risk Prediction in Clinical Practice: The Future Is Now
Kunal N. Karmali and Donald M. Lloyd-Jones

*J Am Heart Assoc.* 2017;6:e006019; originally published April 24, 2017; doi: 10.1161/JAHA.117.006019

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://jaha.ahajournals.org/content/6/4/e006019