Cardiovascular Risk Prediction
Widening the Net

Nina P. Paynter

Accurate assessment of cardiovascular risk in the clinical setting is important for appropriate management and counseling of patients and for identification of biological pathways. Classic models, including the Framingham risk score,[1] have combined strong correlates of risk, such as age, with modifiable factors, such as blood pressure and lipids, into a single model. That structure has been retained by the most recent pooled cohort equations from the American College of Cardiology/American Heart Association guidelines.[2] The article by Ambale-Venkatesh et al[3] in the current issue uses a different framework for prediction and risk factor identification and results in an array of identified risk factors across multiple outcomes.

Article, see p 1092

Ambale-Venkatesh et al build on the strengths of MESA (Multi-Ethnic Study of Atherosclerosis), which collected extensive baseline and outcome data on the participants and included an ethnically diverse cohort by design. The MESA cohort has been a source of key insights in the use of markers, such as coronary calcium to predict risk,[4] and the findings in MESA have been widely replicated in other cohorts. In this article, the authors took advantage of as much of the information as possible, examining 735 variables across a range of different domains, including multiple imaging modalities, laboratory biomarkers, and electrocardiographic measures in addition to more traditional risk factors and demographic measures. They were also able to use multiple end points including heart failure and atrial fibrillation in addition to stroke, coronary heart disease, and death.

The authors used the random forest method to generate prediction models, and their primary result was to rank the variables by their importance in predicting each outcome, optimized in two thirds of the sample and tested in the remaining one third. Random forests, because they explore combinations of many different variables through the generation of the forest of trees, are well suited to identification of factors that are consistently important. Quantifying importance can be done in multiple ways, the most common of which are based on comparisons with a random null. These include the Gini index, which assesses the effect of random assignment at the splits incorporating the variable of interest, and permutation, which runs the model using permuted values for the variable of interest to assess the effect. The method used by the authors instead reports a characteristic of the forest, looking at how close the variable of interest usually is to the root (or first split) of the tree. Stronger predictors are more likely to appear earlier in the tree, where they have the largest effect. As shown in the results, the differences in method lead to some noise in the exact ordering of the factors, but many of the same risk factors are selected by each method. Random forest models also offer an opportunity to identify interactions between variables because of their hierarchical structure, with some suggestion that the Gini method performs best at this.[5]

The variable importance findings from Ambale-Venkatesh et al highlight a consistent set of markers across all outcomes, including blood biomarkers, such as NT-proBNP (N-terminal pro-B-type natriuretic peptide), tumor necrosis factor-α soluble receptor, and interleukin-2 soluble receptor, measures of subclinical disease, such as carotid intima media thickness, coronary artery calcium score and ankle-brachial index, and magnetic resonance imaging and ECG markers. Although the magnetic resonance imaging and ECG markers are not identical across the outcomes, the individual markers may be highly correlated with each other, resulting in generally high prediction across all outcomes rather than being highly specific to a single outcome. Indeed, specificity of risk factors to outcomes, and even models to outcomes, remains unexplored and would be of interest in future work.

The authors also compare the random forest models to forward-selected Cox models and Cox models based on the top variables identified by random forests. Although the random forest models do perform nominally better, the magnitude and significance of the improvement is unclear. Both the nominal improvement and the unclear significance are also true for the comparisons with more traditional risk scores developed both inside and outside of the MESA cohort. Future work is needed to compare the models developed by the authors, ideally in an independent cohort, with other MESA models and with more traditional scores before a confident statement of improvement can be made. Further work to establish which factors identify new pathways or targets for intervention will also be needed.

More importantly for risk prediction, the underlying tradeoffs between a final single model, which is easily implemented and understood, compared with a more complicated black-box model, which is summarized rather than fully described, needs to be considered. In the generation of the Reynolds risk score, a similar method was used in identification of the key variables and model structure, but the goal was to use those insights to generate a well-performing Cox model.[6] This approach resulted in a model that has performed

From the Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA.

Correspondence to Nina P. Paynter, PhD, Division of Preventive Medicine, Brigham and Women's Hospital, 900 Commonwealth Ave E, 3rd Floor, Boston, MA 02215. E-mail npaynter@partners.org

(*Circ Res*. 2017;121:1032-1033.
DOI: 10.1161/CIRCRESAHA.117.311868.)

Circulation Research is available at http://circres.ahajournals.org
DOI: 10.1161/CIRCRESAHA.117.311868

well in other cohorts and shown minimal overfitting.[7] Less research has been done testing the performance of other model types in new cohorts. This will become increasingly important as additional data are available and analyzable, whether from the linking of new domains, such as metabolomics, genetics, or clinical electronic health record data or from more finely grained data for any given marker, such as longitudinal measurements or untargeted information.

Ambale-Venkatesh et al have presented findings which will spark both biological and methodologic interest. They also serve as a thoughtful attempt to navigate the challenges of increasing information. Exciting new methods to work with large datasets, such as deep learning with neural networks, are constantly being developed. However, as we gain more tools for prediction, we also learn more about where those tools fail, as was highlighted by a recent examination of risk prediction performance by neighborhood.[8] Identifying the most useful ways to deploy the new methods and generate across-the-board improvement in the setting of cardiovascular risk requires thoughtful consideration, iteration, and continued grounding of both the results and model descriptions in relation to clinical use and transparency.

Disclosures

None.

References

1. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97:1837–1847.
2. Goff DC Jr, Lloyd-Jones DM, Bennett G, et al; American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. *J Am Coll Cardiol*. 2014;63:2935–2959. doi: 10.1016/j.jacc.2013.11.005.
3. Ambale-Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Guallar E, Bluemke DA, Lima JAC. Cardiovascular event prediction by machine learning: the Multi-Ethnic Study of Atherosclerosis. *Circ Res*. 2017;121:1092–1101. doi: 10.1161/CIRCRESAHA.117.311312.
4. McClelland RL, Jorgensen NW, Budoff M, et al. 10-year coronary heart disease risk prediction using coronary artery calcium and traditional risk factors: derivation in the MESA (multi-ethnic study of atherosclerosis) with validation in the HNR (Heinz Nixdorf recall) study and the DHS (Dallas heart study). *J Am Coll Cardiol*. 2015;66:1643–1653. doi: 10.1016/j.jacc.2015.08.035.
5. Wright MN, Ziegler A, König IR. Do little interactions get lost in dark random forests? *BMC Bioinformatics*. 2016;17:145. doi: 10.1186/s12859-016-0995-8.
6. 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;297:611–619. doi: 10.1001/jama.297.6.611.
7. DeFilippis AP, Young R, Carrubba CJ, McEvoy JW, Budoff MJ, Blumenthal RS, Kronmal RA, McClelland RL, Nasir K, Blaha MJ. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann Intern Med*. 2015;162:266–275. doi: 10.7326/M14-1281.
8. Dalton JE, Perzynski AT, Zidar DA, et al. Accuracy of cardiovascular risk prediction varies by neighborhood socioeconomic position: a retrospective cohort study [published online ahead of print August 29, 2017]. *Ann Intern Med*. 2017. doi: 10.7326/M16-2543. http://dx.doi.org/10.7326/M16-2543.

KEY WORDS: Editorials ■ cardiovascular diseases ■ humans ■ primary prevention ■ risk factors

Circulation Research
JOURNAL OF THE AMERICAN HEART ASSOCIATION

Cardiovascular Risk Prediction: Widening the Net

Nina P. Paynter

Circ Res. 2017;121:1032-1033
doi: 10.1161/CIRCRESAHA.117.311868
Circulation Research is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0009-7330. Online ISSN: 1524-4571