

TITLE PAGE

Title: Predicting the future of cardiovascular risk prediction

Author names and affiliations:

Amitava Banerjee¹

ami.banerjee@ucl.ac.uk

¹Farr Institute of Health Informatics Research, University College London, London

Corresponding author: Dr Amitava Banerjee

Present address: Farr Institute of Health Informatics Research

222 Euston Road

London

UK

NW1 2DA

Tel: 02035495449

Email: ami.banerjee@ucl.ac.uk

Word count: 822 words

Predicting the future of cardiovascular risk prediction

The iconic Framingham Heart Study paved the way for many prospective epidemiologic cohort studies of cardiovascular disease (CVD) around the world, and the subsequent development of tools which can predict risk of future disease. There have been tools for a broad spectrum of outcomes from absolute CVD risk at varying time points to lifetime risk(1) and vascular age(2). Multiple scores for incident CVD have been derived and validated, and adopted in consensus guidelines for use in routine practice.

However, the major burden of CVD events occurs in individuals who have already been diagnosed with recurrent disease, yet even in high-income countries, adherence and persistence to secondary prevention drugs is far from optimal(3). By comparison with incident CVD, relatively few risk prediction tools are available for clinicians to use in secondary prevention. These patients have a wide range of absolute risk for CVD. Moreover, many of the scores for this purpose have been developed in hospital populations rather than primary care, where they would arguably have more utility. In this issue of *Heart*, Poppe and colleagues, report a novel prediction tool (PREDICT-CVD) for secondary CVD prevention, derived using data from 24,927 patients in the New Zealand primary care setting with 1,480 CVD events within 2 years of a baseline CVD risk assessment. Both discrimination and calibration of PREDICT-CVD were very good, although it overestimated risk in the highest-risk individuals, particularly those with heart failure [Poppe et al. *Heart* 2017].

There are several key principles which the PREDICT-CVD score illustrates. First, the authors confirm that locally derived scores are likely to have the best performance in that population. Such scores are more likely to be taken up by clinicians. On the other hand, more universal scores may be required for the purpose of national or cross-national implementation and monitoring. Therefore, it is always important to consider the characteristics of the target population for the risk prediction tool, as well as the clinicians and the context in which the score will be used. Second, simple risk prediction tools derived from routine clinical data have as much clinical utility, if not more, than more complex tools involving biomarkers and genomics. Third, across different CVD types, the use of electronic health records rather than traditional research cohort study designs is increasing and probably have greater application in clinical practice(4).

However, there are four issues which still remain. First, in PREDICT-CVD, a CVD event was defined as MI, ischaemic or haemorrhagic stroke, or CVD death, defined from ICD-10-AM codes. Therefore, the score is concerned with individuals with atherosclerotic disease. Given the increasing importance of heart failure, atrial fibrillation and other comorbidities as well as the growing number of risk prediction tools for different forms of CVD, there may be a role for simplification of tools for clinicians and patients in order to look at the possibility of a composite score for a broader range of CVD event types.

Second, the use of the PREDICT-CVD score and other scores offers the promise of personalisation of drug and lifestyle therapy(5), but, in reality, may not have a substantial impact on outcomes. In this same population in New Zealand, persistence to secondary prevention medications has already been studied. Among patients

untreated at baseline, individuals with prior CVD had the highest dispensing rates for blood-pressure-lowering and lipid-lowering, and incrementally higher dispensing rates were noted as CVD risk group increased. Given that over half of all CVD events occurred in the 20% of the population at highest risk of future CVD, it is reasonable to use the PREDICT-CVD score in these patients to optimise their management. However, the same study also showed that about two-thirds of patients with prior CVD were already using appropriate medications at baseline, and continued to do so after 3 years of follow-up(6). In some subpopulations, such as elderly, there is evidence that due to high absolute risk reductions, treating all patients may be more beneficial than risk prediction-based treatment for secondary CVD prevention(7).

Third, risk prediction is only one stage in the approach to CVD management and should not be considered in isolation. Communication and treatment to mitigate the predicted risk follow from risk prediction. However, the impact of clinical prediction rules and risk prediction tools is rarely evaluated(8).

Finally, at present, the potential of EHR for development and validation of risk prediction tools across diseases remains largely untapped with few multi-centre studies and even fewer with validation across sites(9). Although the PREDICT-CVD initiative uses routine clinical data, the current model uses a derivation cohort which recruited from 2006 to 2012, and a validation cohort which was from 2005 to 2010. The future scenario may well be real-time data analytics to use in a contemporaneous population. In an era of advanced big data analytics, machine learning and artificial intelligence(10) offer new ways of predicting risk. The current paradigm of multi-variable Cox regression models informed by prospective cohort studies testing the incremental gain of novel biomarkers one-by-one needs to be challenged.

References

1. Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation*. 2010;121(15):1768-77.
2. Groenewegen KA, den Ruijter HM, Pasterkamp G, Polak JF, Bots ML, Peters SA. Vascular age to determine cardiovascular disease risk: A systematic review of its concepts, definitions, and clinical applications. *European journal of preventive cardiology*. 2016;23(3):264-74.
3. Yusuf S, Islam S, Chow CK, Rangarajan S, Dagenais G, Diaz R, et al. Use of secondary prevention drugs for cardiovascular disease in the community in high-income, middle-income, and low-income countries (the PURE Study): a prospective epidemiological survey. *Lancet*. 2011;378(9798):1231-43.
4. Goldstein BA, Navar AM, Pencina MJ. Risk Prediction With Electronic Health Records: The Importance of Model Validation and Clinical Context. *JAMA Cardiol*. 2016;1(9):976-7.
5. van der Leeuw J, Ridker PM, van der Graaf Y, Visseren FL. Personalized cardiovascular disease prevention by applying individualized prediction of treatment effects. *Eur Heart J*. 2014;35(13):837-43.
6. Mehta S, Wells S, Grey C, Riddell T, Kerr A, Marshall R, et al. Initiation and maintenance of cardiovascular medications following cardiovascular risk assessment in a large primary care cohort: PREDICT CVD-16. *European journal of preventive cardiology*. 2014;21(2):192-202.
7. Stam-Slob MC, Visseren FL, Wouter Jukema J, van der Graaf Y, Poulter NR, Gupta A, et al. Personalized absolute benefit of statin treatment for primary or secondary prevention of vascular disease in individual elderly patients. *Clin Res Cardiol*. 2017;106(1):58-68.
8. Wallace E, Uijen MJ, Clyne B, Zarabzadeh A, Keogh C, Galvin R, et al. Impact analysis studies of clinical prediction rules relevant to primary care: a systematic review. *BMJ Open*. 2016;6(3):e009957.
9. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JP. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. *J Am Med Inform Assoc*. 2017;24(1):198-208.
10. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J*. 2016.