

# Cardiovascular risk prediction- now and the future

**Ian M. Graham**

Department of Cardiology, Trinity College Dublin, Ireland  
Professor of Cardiovascular Medicine, Trinity College Dublin  
Recent past Chair, European Society of Cardiology Cardiovascular Risk Collaboration  
Project Leader, SCORE risk estimation projects

## ABSTRACT

### Keywords

Cardiovascular risk estimation; lifetime risk



© 2024 The Authors  
Published by SITeCS

Current cardiovascular risk estimation systems that estimate 10-year risk based on cohort studies starting at around age 40 have probably reached their limits based on current methods.

The challenges are to develop new systems that will permit personalised risk estimation earlier in life with better estimates of true lifetime risk and likely treatment benefits. We outline approaches to address these issues.

Disclosures: I have no conflicts of interest, intellectual or financial.

Acknowledgements: This review is based in part upon the work of the European Society of Cardiology's Cardiovascular Risk Collaboration, of which the author was a founder and co-Chair. It is a pleasure to acknowledge Panos Vardas, Adam Timmis, Victor Aboyans, Emanuele di Agelantonio, Dirk De Bacquer, Brian Ference, Frank Visseren, Bill McEvoy and the staff of the ESC's European Heart Health Institute.

Terminology: It is perhaps pedantic to observe that, although the term 'cardiovascular risk prediction' has become embedded in the cardiological literature, what we in fact do is to estimate the risk to allow a prediction of the likelihood of a future clinical event. The term 'screening' strictly applies to testing to assess the likelihood of disease. 'Health risk assessment' is a wider term that includes demographics, social factors, lifestyle and assessment of risk factors.

Received 15 February 2024; accepted 18 March 2024

## Why assess cardiovascular disease (CVD) risk?

Cardiovascular risk assessment is used to guide management decisions. In general, the higher the risk, the more intense will be the preventive efforts required.

In most people, the risk of a future atherosclerotic CVD event is the product of the combined effect of a number of risk factors such as hyperlipidaemia, hypertension, smoking and diabetes. The clinical estimation of the effect of such combinations is unreliable, which is the rationale for risk scoring systems [1].

## When to assess CVD risk

Current risk estimation systems are based on cohort studies that started at about age 40 and so estimate risk from then on. This misses 40 years of exposure to risk, in addition to in-utero risk. The future, discussed below, is clearly to develop systems that can estimate risk much earlier in life.

Risk evaluation may be opportunistic (when a person presents for another reason) or systematic, either population wide or in defined groups with known risk factors such as smoking or diabetes. Population wide risk assessment allows improvement in risk factors but it has been difficult to demonstrate improved outcomes [2], and hence cost-effectiveness is uncertain. Many countries prefer a combination of opportunistic evaluation and evaluation in those with known risk factors.

## How to assess risk

Both the 2021 ESC Prevention Guidelines [2] and The 2019 ESC/EAS Guidelines for the management of dyslipidaemias [3] define categories of risk. In general the latter adopts a simpler approach but both agree that subjects with established CVD have declared themselves to be at very high risk and intensive and immediate risk factor advice is advised.

In *apparently health persons*, The European Society of Cardiology 2021 Guidelines on the Prevention of Cardiovascular Disease in Clinical Practice [2] recommend the use of SCORE2 [4] or, in persons over 70 years, SCORE2-OP [5] for risk assessment. These tools are calibrated for four risk regions of Europe and can be re-calibrated for other countries. HeartScore is a simple, interactive online calculator that facilitates the use of SCORE ([www.heartscore.org](http://www.heartscore.org))

In those at intermediate risk, screening for asymptomatic disease, for example through coronary artery calcium scoring may help to re-classify risk [2].

In America, use of the Pooled Cohort equation is recommended [6], more recently supplemented by the PREVENT calculator [7].

## Limitations of current risk estimation systems

Current risk estimation systems are derived from cohorts studies, most of which started at about age 40, in other words after many years of exposure to risk. Strictly speaking, they apply only to the population from which they were derived. They may work well in other similar populations, or can be re-calibrated for others [4, 5], but the problem remains that the risk estimates apply to groups rather than individuals.

Current techniques such as Cox derive beta-coefficients that are essentially multipliers and cannot easily estimate complex interaction effects within different combinations of risk factors. Further, risk estimates are dominated by the effect of age, especially when risk is expressed over 10 years. Current estimates of lifetime risk also start too late, usually around age 40.

The impact of genetic factors has been underestimated. While polymorphisms affecting risk may have a seemingly small impact on 5-10 year risk, their impact on true lifetime risk, from birth on, may be much greater than is generally appreciated [8].

### Can we see the future?

Ideally, one would like to be able to:

- allow better for the dominance of age in risk estimation
- estimate true lifetime risk from early in life
- approach more individualised estimates of risk
- make more precise estimates of treatment benefits
- explore integrating in-utero determinants of risk

Ference and others [8, 9] have pointed out that Mendelian randomisation studies suggest the impact of polymorphisms on risk has

been greatly underestimated, given that they function from birth on. These effects may be direct or, probably more importantly, through their effect on determining the rate of rise of risk factors such as LDL cholesterol and blood pressure.

This has led to a suggestion to move from 10-year risk to an exposure time model in which risk is expressed as mmol/years of LDL cholesterol, mmHg/years of blood pressure or indeed years of exposure to total risk. Such an approach integrates the rate of rise of risk with time which is likely to parallel the development of atherosclerosis. Thus, given several measurements of risk over several years in younger persons, it should be possible to give a personalised estimate of risk much earlier in life than is currently possible to allow true preventive action early in life. This Mendelian Randomisation-based approach can also permit more precise and logical estimates of likely treatment benefits.

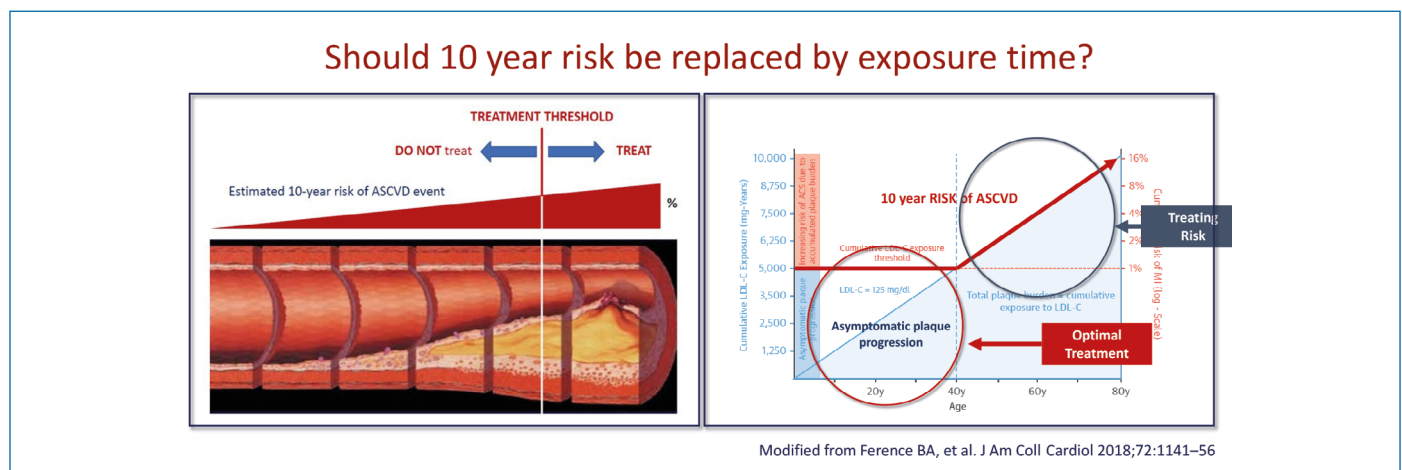
The exposure time approach may be summarised as depicted in **Figure 1**.

Vardas has commented on the transition from ancient medicine through modern medicine to what he terms metaclinical medicine [10]. The latter includes, inter alia, artificial intelligence (A-I) and decision-making models. AI is indeed necessary for the approaches summarised above and the need will grow, necessitating dialogue between medical statistics and A-I [11].

Generative A-I can be used to develop risk estimation systems. Starting with existing large data sets, A-I is used to examine patterns and interactions faster and more efficiently than can be done with conventional statistics. A subset is used for machine learning followed by deep learning such as layered neural networks and generative A-I to produce new content.

Alas, it is of course not that simple [11]. Issues include:

- Data quality. No system can allow for poor quality or non-representative data
- Conclusions based on inadequate data may be re-enforced- the 'self-fulfilling prophecy'
- Arising, results may not seem justified by expectations based on the training set- so-called A-I 'hallucination'
- Conventional statistics use clearly verifiable methods. The deeper one goes into machine learning, the more opaque the process becomes
- 'Data-set shift' [12], in which there is a mis-match between the machine-learning model's training data and the results when the model is applied. This is of course not necessarily a fault of the process if it is applied to a very different population



**Figure 1** | The exposure time model compared with estimation 10 year risk at, say age 40 -modified from the concepts expressed in [9].

Will these advances produce better outcomes? Ference (personal communications and late-breaking session presentations at the European Society of Cardiology and American Heart Association and American conferences) provides compelling arguments. Yet it is hard to envisage how to design a randomised controlled trial to compare usual care with conventional risk estimation and with the A-I based exposure time approach. The clinician is advised to simply see if a risk estimate, based on whatever estimation process, is plausible.

A comparison of machine learning and conventional risk estimation [13] found in favour of machine learning by a modest amount but with substantial caveats-

“In this systematic review and meta-analysis, ML algorithms were found to be superior to traditional risk equations on comparison of C-statistics in the pooled meta-analysis of 11 studies.

However, findings need to be interpreted with caution as the quality of studies was sub-optimal- with all studies performed on retrospective cohorts, half of the studies providing no comparative calibration metrics, and only three with external validation. In addition, most studies were assessed to have a high risk of bias”.

Finally, should determinants of risk in utero [14] be incorporated into a single approach to risk estimation? Those with low birth weight may benefit from early assessment and management of risk. A fully integrated approach to risk from conception through childhood and into adult life would seem logical.

## Conclusion

Risk estimation had become rather static. We now enter an exciting new era of risk estimation based on A-I supported risk estimation, Mendelian randomization and risk expressed as exposure time that has the potential to permit personalised risk estimation early in life with better estimates of likely treatment effects.

## References

- [1] De Backer G. New insights in cardiovascular risk estimation and stratification. *e-Journal of Cardiology Practice* 2022; 22.
- [2] Visseren FLJ, Mach F, Smulders YM, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J* 2021; 42:3227-337.
- [3] Mach F, Baigent C, Catapano AL, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *Eur Heart J* 2020; 41:111-88.
- [4] SCORE2 working group and ESC Cardiovascular risk collaboration. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur Heart J* 2021; 42:2439-54.
- [5] SCORE2-OP working group and ESC Cardiovascular risk collaboration. SCORE2-OP risk prediction algorithms: estimating incident cardiovascular event risk in older persons in four geographical risk regions. *Eur Heart J* 2021; 42:2455-67.
- [6] Arnett DK, Blumenthal RS, Albert MA, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* 2019; 140:e596-e646.
- [7] Khan SS, Matsushita K, Sang Y, et al. Development and Validation of the American Heart Association's PREVENT Equations. *Circulation* 2024; 149:430-49.
- [8] Ference BA, Ginsberg HN, Graham I, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur Heart J* 2017; 38:2459-72.
- [9] Ference BA, Graham I, Tokgozoglu L, Catapano AL. Impact of Lipids on Cardiovascular Health: JACC Health Promotion Series. *J Am Coll Cardiol* 2018; 72:1141-56.
- [10] Vardas PE, Vardas EP, Tzeis S. Medicine at the dawn of the metaclinical era. *Eur Heart J* 2023; 44:4729-30.
- [11] Hunter DJ, Holmes C. Where Medical Statistics Meets Artificial Intelligence. *N Engl J Med* 2023; 389:1211-9.
- [12] Lea AS, Jones DS. Mind the Gap - Machine Learning, Dataset Shift, and History in the Age of Clinical Algorithms. *N Engl J Med* 2024; 390:293-5.
- [13] Liu W, Laranjo L, Klimis H, et al. Machine-learning versus traditional approaches for atherosclerotic cardiovascular risk prognostication in primary prevention cohorts: a systematic review and meta-analysis. *Eur Heart J Qual Care Clin Outcomes* 2023; 9:310-22.
- [14] Ardisino M, Morley AP, Slob EAW, et al. Birth weight influences cardiac structure, function, and disease risk: evidence of a causal association. *Eur Heart J* 2024; 45:443-54.