Risks in estimating risk

Ian M. Graham* and Marie-Therese Cooney

Trinity College, Dublin, Ireland

Online publish-ahead-of-print 7 November 2013

This editorial refers to ‘SCORE performance in Central and Eastern Europe and former Soviet Union: MONICA and HAPIEE results’, by O. Vikhireva et al., on page 571

‘Prediction is very difficult, especially about the future’
(attributed to Niels Bohr)

‘I told you I was ill’
(Spike Milligan’s epitaph)

The paper by Vikhireva and colleagues is a timely reminder that one size does not fit all when estimating the risk of cardiovascular death. The authors examined the performance of the European Society of Cardiology (ESC) cardiovascular disease (CVD) risk estimation system SCORE in the Czech Republic, Poland, Lithuania, and Russia, using data from the mid 1980s (MONICA) and early 2000s (HAPIEE). Given acknowledged methodological limitations, the high-risk version of the SCORE chart estimated risk fairly well in the older cohorts, apart from underestimating it substantially in Russia. In the more recent cohorts, SCORE overestimated risk apart from in Russia. It was concluded that the low-risk version of SCORE might now be more appropriate for the Czech Republic and Poland. The fact that re-calibrated versions exist for these two countries was not discussed because it was felt that there was insufficient information on the methods used; it is not clear whether this information was sought.

The estimation of CVD risk is not an exact science. The frequently used Cox’s proportional hazard model uses regression coefficients to estimate risks relative to an absolute baseline risk. The coefficients are assumed to remain constant over time and in the context of different combinations of other risk factors. Explanatory variables are considered to act multiplicatively on the hazard function. At best, these assumptions may be regarded as usable approximations to ‘truth’. For example, different combinations of risk factors may interact in complex ways that are difficult to model. We have recently shown that beta coefficients are by no means constant as a person ages.

Cardiovascular disease mortality is less stable than one might imagine. It can change rapidly over time as countries undergo the epidemiological transition of eating and smoking because they can afford to, until they appear to reach a sort of saturation point beyond which risk tends to decline. Figure 1 demonstrates the changes in age-standardized coronary heart disease (CHD) mortality rates from 1970 to 2006 in men from the countries studied by Vikhireva et al. and average rates in Scandinavian, Southern and Eastern European countries for comparison, calculated from the World Health Association (WHO) mortality statistics. These time trends, given fluctuations in the Russian data, may explain most if not all of the findings of Vikhireva et al. They generally relate rather more to changes in lifestyle than to improvements in treatments.

Some persons announce that they are at high risk of cardiovascular events by virtue of having overt CVD, diabetes, or renal impairment, or an extremely high single risk factor. However, for most seemingly healthy people, risk is the product of several factors. Risk estimation systems attempt to estimate the combined effects of several risk factors. These considerations are summarized in Table 1, which shows the risk categories defined by the European Guidelines on CVD prevention. As noted, such systems are not capable of analysing interaction effects in any detail so the estimates will necessarily be approximate.

After a risk estimation system is derived, it is inevitable that it will overestimate risk in countries in which CVD mortality has declined and underestimate it if mortality has increased. As noted above, this is likely to explain a substantive part of the findings of Vikhireva and colleagues.

So what are the options in trying to improve risk estimation? The least likely to be effective is to hope to find a new mega-risk factor—the literature is littered with failed attempts, not least our own work on homocysteine and risk. It is improbable that any other factor will fulfil the criteria for causality as well as do hyperlipidaemia, smoking, and hypertension. As for that hoped-for portfolio of polymorphism that would solve all, dream on...

The challenge is to use the tools that we have appropriately rather than to expect major refinements in an inexact science. Nevertheless, use of different risk calculators may result in substantial variations in risk estimates. Jackson et al. have pointed out two possible reasons for this—the use of different CVD outcomes, and failure to recalibrate the risk system to the population to which it is applied.
Are risk estimation systems such as Framingham or SCORE widely applicable? When deciding what is best for one’s own country, there are several options.

If the country is recognizably similar to the population from which the risk chart has been derived in terms of baseline CVD mortality and risk factor prevalences, it may well be appropriate to use the chart as it is. The current ESC Guidelines on CVD prevention offer guidance on this.

Secondly, the risk chart may be re-calibrated to allow for secular changes in mortality and risk factor prevalences. This process appears to perform fairly well and, with regard to SCORE, has been undertaken for 10 European countries.

Thirdly, the country may have its own cohort data permitting a risk estimation system that is likely to be directly applicable to that population to be derived. Eventhen, within-population heterogeneity, for example in the USA, may limit its universal applicability.

More sophisticated approaches may be feasible in countries with appropriate data systems. In the UK, QRISK uses general practitioner records of millions of patients to update the risk estimates. Issues regarding representativeness and missing data may recede as data accumulate. In New Zealand, the PREDICT web-based decision support system generates individualized risk estimates and personalized treatment recommendations.

Finally, the use of risk age may circumvent some of these problems. Specialists in prevention and public health generally prefer absolute to relative risk in advising on risk management, and this is sensible—a large relative risk of, say, 10 is not very important if the absolute risk is 0.0001%. Yet it can help in advising young persons with a low absolute risk that this will rise over time if the relative risk is high. Risk age is a way of expressing relative risk that may be meaningful to patients. For example, a 40 year old with a risk of 3% because of multiple risk factors has the same risk as a 60 year old with no risk factors, indicating a risk age of 60. Importantly, and not well recognized, risk age appears to be independent of baseline risk or changes in baseline risk over time, and so retains its clinical utility without requiring re-calibration.

The vulnerability of the performance of any risk estimation system to secular trends in CVD mortality poses a challenge to Guideline writers and to all trying to make prevention accessible and easy to understand. In the 2012 European Guidelines on CVD prevention, 25 countries were classified as low risk, compared with eight in the 2007 Guidelines. One tries to balance accuracy with the need to avoid confusion caused by too many.

Figure 1: Age-standardized coronary heart disease mortality rates in men aged under 65 years in the countries studied and European regions, from the World Health Organization mortality statistics. South European countries, average of France, Greece, Italy, and Spain; Eastern European countries, average of Poland, Hungary, Romania, and Bulgaria; Scandinavian countries, average of Denmark, Finland, Norway, Sweden, and Iceland.
changes in advice. One can conceive of a system of continuous electronic re-calibration of risk estimates, but this is beyond the capability of current data systems.

Some of the methodological issues in the work of Vikhireva et al. 1 may be worthy of further discussion in another forum. The authors calculated the C-statistic for the discrimination of SCORE when dichotomized using a cut-off point of 5% 10-year risk of fatal CVD. It is likely that if the discrimination of SCORE as a continuous variable for estimating risk of future CVD was examined the C-statistic would have been considerably higher. This may be more appropriate since risk is a continuum, as acknowledged in the guidelines. 9,31 Further, some of the the results, such as poor calibration in some of the more recent cohorts, are not unexpected. It would be impossible for a system to predict risk well in two cohorts from the same country but two decades apart when CVD mortality rates have changed substantially over the time frame.

In conclusion, the study of Vikhireva and colleagues is helpful in confirming that risk estimation systems must be used judiciously in countries where baseline risk is substantially different from that of the SCORE cohorts. We have outlined approaches to this problem. Resources permitting, increased use of the electronic version of HeartScore12 with as frequent recalibrations as is feasible, flagged by appropriate advisory notes, may be preferable to slavish adherence to paper charts that are more difficult to update. HeartScore already generates personalized patient management advice. There is a compelling need for ongoing cohort data collection throughout Europe and for more sophisticated data linkage systems to refine and facilitate risk estimation and management.

Conflict of interest: none declared.

References

**Table 1** Risk categories defined by the European guidelines on cardiovascular disease prevention

<table>
<thead>
<tr>
<th>Risk category</th>
<th>Subjects with any of the following:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Very high Risk:</strong></td>
<td>• CVD</td>
</tr>
<tr>
<td></td>
<td>• Type 2 diabetes, or type 1 diabetes &amp; target organ damage</td>
</tr>
<tr>
<td></td>
<td>• Patients with moderate to severe CKD (GFR &lt;60mL/min/1.73m²)</td>
</tr>
<tr>
<td></td>
<td>• SCORE ≥ 10%</td>
</tr>
<tr>
<td><strong>High Risk:</strong></td>
<td>• Markedly elevated single risk factors such as:</td>
</tr>
<tr>
<td></td>
<td>- Familial dyslipidaemias</td>
</tr>
<tr>
<td></td>
<td>- Severe hypertension</td>
</tr>
<tr>
<td></td>
<td>• SCORE ≥ 5% and &lt;10%</td>
</tr>
<tr>
<td><strong>Moderate Risk:</strong></td>
<td>• SCORE ≥ 1% and &lt;5% at 10 years, further modulated by:</td>
</tr>
<tr>
<td></td>
<td>• family history of premature CAD</td>
</tr>
<tr>
<td></td>
<td>• abdominal obesity</td>
</tr>
<tr>
<td></td>
<td>• physical activity pattern</td>
</tr>
<tr>
<td></td>
<td>• HDL-C</td>
</tr>
<tr>
<td></td>
<td>• TG</td>
</tr>
<tr>
<td></td>
<td>• hsCRP</td>
</tr>
<tr>
<td></td>
<td>• social class</td>
</tr>
<tr>
<td><strong>Low Risk:</strong></td>
<td>• SCORE less than 1% and free of qualifiers</td>
</tr>
</tbody>
</table>

References