EDITORS’ NETWORK

Almanac 2012: cardiovascular risk scores. The national society journals present selected research that has driven recent advances in clinical cardiology

Almanac 2012: Scores de risco cardiovascular. As revistas das sociedades cardiológicas nacionais apresentam investigação selecionada que levou a avanços recentes na cardiologia clínica

Jill P. Pell

Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

Received 31 October 2012; accepted 31 October 2012

Global risk scores use individual level information on non-modifiable risk factors (such as age, sex, ethnicity and family history) and modifiable risk factors (such as smoking status and blood pressure) to predict an individual’s absolute risk of an adverse event over a specified period of time in the future. Cardiovascular risk scores have two major uses in practice. First, they can be used to dichotomise people into a group whose baseline risk, and therefore potential absolute benefit, is sufficiently high to justify the costs and risks associated with an intervention (whether treatment or prevention) and a group with a lower absolute risk to whom the intervention is usually denied. Second, they can be used to assess the effectiveness of an intervention (such as smoking cessation or antihypertensive treatment) at reducing an individual’s risk of future adverse events. In this context, they can be helpful in informing patients, motivating them to change their lifestyle, and reinforcing the importance of continued compliance.

How have risk scores evolved?

Our understanding of how best to measure and respond to risk has evolved over a number of years. Historically, individual risk factors were measured and managed in isolation, but this has been replaced by the adoption of global risk scores that calculate overall risk based on a range of risk factors. Also, the opportunistic use of risk scores among people who present to healthcare workers has been replaced by increased use of either mass screening or targeted screening of at-risk populations in an effort to identify unmet need and reduce health inequalities. The integration of risk calculators into administrative software packages and online access have made risk scores readily accessible to all general practitioners in the UK. The scope of risk scores has recently widened beyond coronary heart disease to other conditions, such as heart failure and diabetes mellitus. Also, as new biomarkers for cardiovascular disease have been identified, there has been an increasing number of studies examining whether they can add value to existing risk scores. Finally, as investigators have identified genetic loci associated with cardiovascular conditions, studies have started to address...
whether they could play a role in risk prediction, either in isolation or combined with traditional risk factors.

Our approach to evaluating the performance of risk scores has also evolved over time. Initially, methods were adopted from the assessment of screening tests, using measures of discrimination such as sensitivity and specificity. As many predictive models could be expressed as continuous variables, interest grew in assessing the performance of predictive models across the whole range of values. This was achieved by plotting sensitivity versus 1-specificity for all values to produce a receiver operating characteristic (ROC) curve. The area under the ROC curve, also referred to as the c statistic, ranges from 0.5 (no predictive ability) to 1.0 (perfect discrimination). For use in clinical or public health practice, a continuous measure of risk needs to be reduced to two or more categories, but the ROC plot can be useful in determining the best cut-off values to apply. More recently, investigators have used reclassification between different risk groups to compare the discriminatory performance of different risk scores. Results can be presented simply as the total percentage of patients into a different risk group, but the preferred measure is the net reclassification index, which is calculated from: (proportion of cases moving up — proportion of cases moving down) — (proportion of controls moving up — proportion of controls moving down).

One hundred and ten ways to measure risk!

Historically, cardiovascular risk scores have focused on coronary heart disease; either predicting the risk of adverse events in the general population or among patients with established disease such as those presenting with acute coronary syndromes. There are now 110 different cardiovascular risk scores that have been developed for use in the general population.2 More recent risk scores, such as ASSIGN (ASsessing cardiovascular risk using SIGN) and QRISK (QRESEARCH cardiovascular risk algorithm), have differed from earlier scores by incorporating socioeconomic deprivation and family history into the measurement of global risk.3-5 As a result, they have been able to overcome some of the limitations of earlier risk scores, which tended to introduce socioeconomic bias into the detection and treatment of cardiovascular risk.4 However, the performance of all risk scores is dependent on ready access to complete and accurate data. In a recent study, in which they applied six risk scores to routine general practice data, de la Iglesia and colleagues4 highlighted missing data as a concern, especially in relation to family history.

Knowledge of risk scores can translate into improved prescribing and reduced risk.6 However, in a recent systematic review, Llew and colleagues7 highlighted a number of problems in the development of risk scores including a lack of standardisation in the measurement of risk predictors and outcomes, and failure of most studies constructing new risk scores to take account of individuals who are already taking medications that modify risk measurement, such as antihypertensive and lipid-lowering agents. The latter may be misleading because primary prevention should, ideally, be directed at individuals before the development of risk factors and the occurrence of premature disease. One of the limitations of existing risk scores based on events over a fixed period of time, commonly 10 years, is that the score is heavily influenced by age. Therefore, young individuals are unlikely to reach the threshold for intervention irrespective of their current and future risk factors. One approach to identifying the subgroup of young people at increased risk is to use lifetime risk rather than risk over a fixed period. Hippslisly-Cox and colleagues8 recently compared the use of QRisk2 reported as the lifetime risk of cardiovascular disease (in terms of age-sex specific centiles) with it reported as risk over a 10-year period. The former identified a greater a proportion of younger individuals as being at risk of future events. It also classified a greater proportion of individuals from ethnic minority groups and with a positive family history as being at risk of future cardiovascular events. Both factors are associated with an increased risk of premature cardiovascular events. While early identification and prevention are the ideal, the unselected screening of a younger population may, nonetheless, be less cost-effective.

The application of risk scores to patients presenting with acute coronary syndrome is now well established in both research and clinical practice. In a recent Education in Heart paper, Bueno and Fernandez-Aviles9 reviewed 11 risk scores developed for the prediction of adverse events following acute coronary syndrome. Of these, the GRACE (Global Registry of Acute Coronary Events) and TIMI (Thrombolysis in Myocardial Infarction) risk scores have been most widely adopted. Fox and colleagues10 recently reviewed the extent to which the GRACE risk score has been validated and adopted since first developed in 2003. To date, the GRACE risk score has been externally validated in 67 individual studies comprising at least 500 patients with acute coronary syndrome, ST-segment elevation myocardial infarction or non-ST-segment elevation myocardial infarction. The risk score is easy to use in a clinical setting and performs well when compared with other risk scores. Therefore, it has been incorporated into many guidelines including those produced by the European Society of Cardiology, American College of Cardiologists, American Heart Association, Scottish Intercollegiate Guidelines Network and National Institute for Health and Clinical Excellence.

Where next for risk scores?

Attention is now focusing on expanding the use of risk scores beyond coronary heart disease. Two recent studies have developed risk scores for use in patients with heart failure. The HF-Action (Heart Failure: A Controlled Trial Investigating Outcomes of Exercise TraiNing) risk score was developed using a cohort of patients with chronic heart failure and systolic dysfunction.11 The risk score was derived from information on exercise duration, serum urea nitrogen, body mass index and sex, and performed well at predicting all-cause death within 1-year of follow-up. Nineteen per cent of patients in the top decile for risk score died, compared with 2% in the bottom decile. The score had a c statistic of 0.73. The GWTG-HF (Get With The Guidelines—Heart Failure) risk score was developed using a cohort of patients hospitalised with heart failure.12 The component factors included age, systolic blood pressure, blood urea nitrogen, heart rate, sodium, concomitant chronic obstructive pulmonary disease and race. The risk of in-hospital death ranged from 0.4% to
9.7% across the risk score deciles and performed well among both patients with preserved and impaired left ventricular systolic function with a c statistic of 0.75 in both groups.

Due to the rising prevalence of type II diabetes, there has been increased awareness of the need to target screening and prevention efforts at people with this condition. Van Dieren et al. undertook a systematic review of studies published between 1966 and 2011 that had developed cardiovascular risk scores suitable for use in patients with type II diabetes mellitus. Of the 45 scores identified, only 12 were originally constructed from a cohort of individuals with diabetes and only two of these were restricted to patients in whom diabetes had been recently diagnosed. Only nine studies reported the c statistic. Six scores had undergone internal validation, using bootstrapping or a split sample, and six had been subject to external validation. Two studies had neither internal nor external validation. The authors identified an additional 33 scores that were constructed from the general population but included diabetes as a predictive factor. Only 12 had internally validated their risk score using a split sample, cross-validation or bootstrapping, and only eight had been externally validated in a population with diabetes. Given the increasing prevalence of type II diabetes and its increasing contribution to cardiovascular disease, further research is required in this area.

**Do biomarkers add value?**

Several recently published studies have examined whether the addition of biomarkers improved the performance of risk scores in the general population. A common focus of these studies has been trying to achieve better discrimination within the subgroup of individuals currently classified as having intermediate risk (10-20% risk of an adverse event over 10 years). Melander and colleagues evaluated the added value of a panel of biomarkers, C-reactive protein (CRP), cystatin C, lipoprotein-associated phospholipase A2 (Lp-PLA2), mid-regional pro-adrenomedullin (MR-proADM), mid-regional pro-atrial natriuretic peptide and N-terminal pro-B-type natriuretic peptide (NT-proBNP), in predicting incident cardiovascular events in a Swedish population cohort. There was a non-significant increase in the c statistic. In relation to predicting cardiovascular events, 8% were reclassified overall but only 1% were moved into the high-risk category. There was no net reclassification. Among the intermediate risk group, the addition of biomarkers resulted in reclassification of 16% in terms of their risk of cardiovascular events, but only 3% were moved into the high-risk group. The net reclassification improvement was 7.4%. Therefore, the improvements in classification were largely achieved by down-grading, rather than identifying a greater proportion of high-risk individuals.

Rana and colleagues examined the added value of a series of individual biomarkers in the UK population in predicting coronary events: CRP, myeloperoxidase, paraoxonase, group IIa secretary phospholipase A2, Lp-PLA2, fibrinogen, macrophage chemotactic protein 1 and adiponectin. Reclassification was greatest for CRP, the addition of which resulted in 12% net reclassification improvement overall and 28% in the intermediate group. Zethelius and colleagues examined the added value of four biomarkers (troponin I, NT-proBNP, cystatin C and CRP) when applied to a population cohort of elderly Swedish men. The addition of all four biomarkers significantly increased the c statistic from 0.66 to 0.77. They reported a 26% net improvement in reclassification overall. The studies to date suggest that biomarker assays may improve discrimination when added to existing risk scores. However, their use has cost and logistical implications, particularly if risk scores are applied on a wide scale. Further research is needed on the cost-effectiveness of adding biomarkers to existing risk scores, particularly in relation to general population screening.

Lorgis and colleagues demonstrated that adding NT-proBNP to the GRACE risk score can improve its prognostic value among patients presenting with acute coronary syndrome. Patients with both a high GRACE risk score and high NT-proBNP level had a 50% risk of dying within 1 year of follow-up. This was sixfold higher than the referent group. NT-proBNP was found to be a useful addition across all age groups but particularly in obese patients, in whom NT-proBNP levels were much lower. Similar findings were reported when troponin and brain natriuretic peptide were used in addition to the TIMI risk score. Their addition produced only a slight increase in the c statistic but, as with NT-proBNP, they were able to identify a subgroup of the TIMI high-risk group who were at very high risk of adverse events, and in whom an aggressive approach to drug therapy and interventions might be warranted. Damman and colleagues examined a cohort of patients undergoing primary percutaneous coronary intervention (PCI) for ST-segment elevation myocardial infarction. They demonstrated that the addition of biomarkers (glucose, NT-proBNP and estimated glomerular filtration rate) improved the prediction of mortality, resulting in significant improvements in net reclassification (49%, p < 0.001) and integrated discrimination (3%, p < 0.01).

Risk scores, such as CHADS2-VASC2, can predict the risk of cerebrovascular events among patients with atrial fibrillation, and are used to inform clinical decisions on the use of anticoagulant therapy. A number of biomarkers has now been identified that are associated with the incidence and prognosis of atrial fibrillation. In a recent review paper, Brugts and colleagues highlighted the need for further research to determine whether the use of these biomarkers may improve the existing risk scores and whether they offer the potential for risk prediction at an earlier stage by identifying patients at risk of developing atrial fibrillation or at risk of progressing from the subclinical to permanent stage of the condition.

Many pathophysiological mechanisms contribute to the development of heart failure. Avellino and colleagues reviewed recently identified biomarkers associated with the relevant pathways. They concluded that the biomarkers currently showing most promise, in terms of risk stratification, were Lp-PLA2 (inflammation), neutrophil gelatinase-associated lipocalin and cystatin C (both renal stress), procollagen-1-polypeptide (extra-cellular matrix remodelling), brain natriuretic peptide, NT-proBNP, MR-proADM, soluble ST2 receptor and copeptin (all cardiac myocyte stress), and endothelin 1 (neurohormone regulation). Gustav Smith and colleagues demonstrated that, in terms of predicting incident heart failure and atrial fibrillation in a general population cohort, the addition
of a panel of biomarkers (mid-regional pro-atrial natriuretic peptide, NT-proBNP, MR-proADM, cystatin C, CRP and copeptin) to conventional risk factors improved discrimination. The net reclassification improvement was 22% for heart failure and 7% for atrial fibrillation. Reclassification was mainly achieved by the identification of additional high-risk individuals. In a recent review, Ketchum and Levy suggested that risk scores had an increasing role to play among patients with advanced heart failure whose survival has improved due to therapeutic and technological advances. They suggested that risk scores could be used to assist the selection of patients for transplant-plantation, left ventricle assist devices and implantable cardioverter defibrillators. Haines and colleagues recently developed a risk score to predict post-procedural complications associated with the implantation of cardioverter defibrillators. The risk score was based on 10 readily available variables: age, sex, New York Heart Association class, presence of atrial fibrillation, previous valve surgery, chronic lung disease, blood urea nitrogen, re-implantation for reasons other than battery change, use of a dual chamber or biventricular device and a non-elective procedure. The 4% of the population in the highest risk category possessed a 8% risk of complications, compared with less than 1% in the lowest risk group.

Studies have recently started to address whether non-invasive imaging of the coronary vessels could add value to existing risk scores. The coronary artery calcium score is a marker of vascular injury and correlates well with the overall atherosclerotic burden. Coronary CT angiography can detect non-calcified plaque and indicates the severity of coronary artery stenoses. Both have been shown to be of incremental value in risk prediction among asymptomatic patients, but studies are generally lacking on the utility of incorporating them into risk scores for use among asymptomatic people. Carotid intima-media thickness is a significant predictor of the risk of cardiovascular events in individuals without carotid plaques. When combined with information on the number of segments with plaque, to produce a total burden of carotid atherosclerosis score, the c statistic and net reclassification index are improved by 6.0% and 17.1%, respectively. The cost of imaging is generally greater than for blood biomarkers. Therefore, the incremental cost is likely to be prohibitive in terms of the routine addition to general population risk scores. Cost-effectiveness studies are required to explore whether the additional costs can be justified in a subgroup of asymptomatic individuals identified by existing risk scores.

One of the few studies to assess the cost-effectiveness of adding biomarkers to clinical risk scores examined patients with stable angina who were on the waiting list for coronary artery bypass grafting. They compared the status quo strategy of no formalised prioritisation with prioritisation using a clinical risk score in isolation and prioritisation after supplementing the clinical risk scores with additional biomarker information using a routinely assessed biomarker (estimated glomerular filtration rate), a novel biomarker (CRP), or both. They demonstrated that the addition of the routinely assessed biomarker improved cost-effectiveness in terms of the net effect on lifetime costs and quality-adjusted life-years. In contrast, addition of the novel biomarker was not cost-effective.

Do genetic markers add value?

Cardiovascular disease is a complex condition, with several intermediate phenotypes, to which both environmental and genetic risk factors predispose. As increasing numbers of genetic markers have been identified, it has become increasingly clear that the genetic component is also complex, with relatively small contributions from a large number of genes. Therefore, attention has focused on the development of a multilocus genetic risk score that summates the overall risk from known genetic markers. In the past couple of years, several studies have investigated whether a genetic risk score can add value to established risk scores, some of which already include information on family history. The studies have been undertaken in a variety of populations but have reached consistent conclusions.

Ripatti and colleagues studied seven cohorts of middle-aged men and women recruited from the general populations in Finland and Sweden. They used published studies to identify 13 recently discovered single nucleotide polymorphisms (SNP) associated with either myocardial infarction or coronary heart disease. They constructed a multilocus genetic risk score for each individual by summing the number of risk alleles for each of the 13 SNP weighted by effect size. The genetic risk score was an independent predictor of incident coronary heart disease, cardiovascular disease and myocardial infarction when adjusted for age, sex and traditional risk factors. In comparison with the lowest quintile of genetic risk score, individuals in the top quintile had an adjusted RR of coronary heart disease of 1.66 (95% CI 1.35 to 2.04). However, addition of the genetic risk score to traditional risk factors did not significantly improve the c statistic. There was a significant improvement in net reclassification of people at intermediate risk (10-year predicted risk of 10-20%) but there was no significant improvement in net reclassification overall.

Paynter and colleagues undertook a similar study using a cohort of white professional women in the USA. They used an online catalogue of genome-wide association studies to identify 101 SNP shown to be associated with any form of cardiovascular disease (including stroke) or any intermediate phenotype (such as diabetes and hypertension), and derived a genetic risk score from the sum of all risk alleles without weighting. They also re-ran the analyses including only the 12 SNP shown to be associated with cardiovascular disease. In comparison with the lowest tertile of genetic risk score, individuals in the highest tertile had a higher RR of cardiovascular events (RR 1.22, 95% CI 1.02 to 1.45) but the difference in the absolute 10-year risk of cardiovascular disease in the top and bottom tertiles was small (3.7% vs 3.0%). Unlike family history (which encompasses overall inherited risk), the genetic risk score was not significantly associated with cardiovascular events after adjustment for traditional risk factors. Addition of the genetic risk score produced no significant improvement in either the c statistic or net reclassification.

Qi and colleagues undertook a case-control study of myocardial infarction survivors in Costa Rica. They examined SNP associated with myocardial infarction and coronary artery disease in at least two previous genome-wide association studies. Of the 14 SNP identified from the literature, seven had significant associations with the risk of myocardial
infarction in their Hispanic cohort. These were used to calculate a genetic risk score based on the sum of the risk alleles. They demonstrated a dose relationship, whereby the risk of myocardial infarction increased with increasing genetic risk score and persisted after adjustment for traditional risk factors, including family history. However, addition of the genetic risk score only increased the c statistic from 0.67 to 0.68.

In common with the previous study by Paynter and colleagues,\textsuperscript{30} Thanassoulis and colleagues\textsuperscript{32} calculated two different genetic risk scores: a more restrictive score derived from 13 SNP previously associated with coronary heart disease or myocardial infarction, and a less restrictive score that included an additional 89 SNP associated with intermediate phenotypes. In both approaches, they also used both a simple and weighted count of risk alleles. Finally, they re-ran the restrictive score adding an additional 16 recently identified SNP. The genetic risk scores were applied to the Framingham Offspring Cohort. The restrictive genetic risk score performed better than the less restrictive score and was an independent predictor of both coronary heart disease and cardiovascular events. Nonetheless, it did not improve discrimination or classification even after addition of the additional SNP.

These studies consistently demonstrate that, even if genotypic information is summarised into an overall risk score, it does not improve the performance of existing risk scores and therefore has no obvious clinical utility, at present, in selecting middle-aged people for interventions. Further research is required to explore whether genetic risk scores have any role to play in identifying the subgroup of young people who are most likely to acquire a high-risk score in the future and, if so, the costs, risks and benefits of providing preventive interventions, such as education, to this subgroup at an earlier stage.

**Procedure risk scores**

Faroq and colleagues\textsuperscript{33,34} recently reviewed the use of risk scores for patients undergoing coronary revascularisation. Clinical risk scores, such as PARSONNET (Predictive score for acquired adult heart surgery: Additive and Logistic Regression models) and EuroSCORE (European System for Cardiac Operative Risk Evaluation), have been widely adopted into clinical practice for patients undergoing coronary revascularisation. Anatomy-based risk scores, which contain no clinical information, have been developed using information derived from diagnostic angiography. As coronary artery grafts are used to bypass stenoses and the anastomoses are positioned distal to the diseased segment, additional anatomical information does not significantly improve the performance of clinical risk scores among patients being managed surgically. In contrast, the severity, length and distribution of stenoses are critical to the selection and outcome of patients undergoing PCI. Anatomy-based scores, such as SYNTAX (SYnergy between PCI with TAXus and surgery), have been shown to be predictive of clinical outcomes following PCI,\textsuperscript{35} but visual interpretation of coronary angiograms is subject to interobserver variation. Therefore, functional anatomy-based scores, which incorporate objective information from fractional flow reserve or quantitative coronary angiography, have better prognostic ability.

More recently, a number of risk scores has been developed that combine clinical and anatomical information.\textsuperscript{36-42} The EuroHeart score is constructed from 12 clinical characteristics and four lesion characteristics. It was developed and validated on the 46,064 patients recruited to the EuroHeart Survey of PCI and performed well at identifying patients at risk of in-hospital death, producing a c statistic of 0.90.\textsuperscript{36} The Clinical SYNTAX Score (CSS) combines the anatomically derived SYNTAX score with a modified version of the clinical ACEF (Age, Creatinine and Ejection Fraction) score. Patients in the highest tertile of CSS had higher rates of repeat revascularisation (21%) and major adverse cardiac and cerebrovascular events (MACCE) (32%) over 1-year following PCI, with evidence of a dose relationship across the tertiles.\textsuperscript{37} The CSS had a higher c statistic than either the SYNTAX score or ACEF score used in isolation in relation to predicting both MACCE and all-cause death.\textsuperscript{37} Capodanno and colleagues\textsuperscript{38} compared two combined clinical/anatomical risk scores (the Global Risk Classification and the Clinical SYNTAX risk score), two clinical risk scores (ACEF and EuroSCORE) and one anatomy-based risk score (SYNTAX) among patients with left main stem stenosis undergoing either PCI or coronary artery bypass grafting. The best predictive characteristics were obtained using a clinical risk score (ACEF) for surgical patients compared with a combined clinical/anatomical risk score (GRC) for PCI.

Similarly, Chen and colleagues\textsuperscript{39} compared the combined clinical/anatomical NERS (New Risk Stratification Score) with the CSS in terms of predicting the risk of MACCE over 6 months follow-up, among patients in whom coronary stents were implanted for left main stem stenoses. In comparison with the clinical risk score, the combined score had both higher sensitivity and higher specificity.\textsuperscript{39} Chakravarty and colleagues\textsuperscript{40} also examined patients treated by surgery or PCI for left main stem disease. They compared the performance of a combined risk score, produced by combining the PARSONNET and SYNTAX risk scores, with using the latter, an anatomical risk score, in isolation. Patients were followed up for a median of 3 years. The study suggested that using anatomical information in isolation did not predict outcome following surgery. In contrast, the SYNTAX risk score was predictive among patients undergoing PCI but could be improved by the addition of clinical information.

Many of the risk scores developed for use in patients undergoing coronary revascularisation predicted the widespread adoption of drug-eluting stents and, therefore, perform less well in these patients than in those undergoing balloon angioplasty. Stolker and colleagues\textsuperscript{43} recently developed and validated a risk score that combined clinical, procedural and anatomical information using the EVENT (Evaluation of Drug Eluting Stents and Ischaemic Events) Registry, and evaluated its ability to predict target lesion revascularisation at 1-year follow-up. The relatively simple score was composed of only six variables: age, previous PCI, left main PCI, saphenous vein graft location, minimum stent diameter and total stent length. The investigators demonstrated a threefold difference in target lesion revascularisation between the highest risk and lowest risk categories (7.5% vs 2.2%).
Conclusion

Cardiovascular risk scores have existed for many years but they are still subject to new and interesting research. They are increasingly being applied to conditions other than coronary heart disease, such as type II diabetes and heart failure, which are of increasing importance for public health. New biomarkers have been identified that improve discrimination but, inevitably, the marginal benefit decreases with each additional predictor. Also, improved discrimination needs to be weighed against increased cost and complexity, especially when risk scores are applied to the general population. As highlighted in a recent Heart editorial, ease of use has a major impact on the implementation of risk scores.^{3} Recent research has focused on identifying new biomarkers and evaluating their effectiveness, but there is a paucity of applied research on cost-effectiveness and coverage. This needs to be addressed. The conclusions may differ depending on the location in which risk scores are being measured and the subgroup of the population to which they are applied. To date, there is no evidence that genetic markers improve risk prediction when used in middle-aged populations. If they have a role to play, it may be in younger people in whom traditional risk scores are of little value. Another approach to identifying at-risk individuals at a younger age is lifetime risk. Irrespective of the approach adopted, the cost-effectiveness of earlier screening and intervention needs to be properly evaluated.

Competing interests None.

Provenance and peer review Commissioned; externally peer reviewed.

Heart 2012; -:1e6. doi:10.1136/heartjnl-2012-302143

References