Cardiovascular disease risk estimation: Why, when and how?

Hermione Price

Risk assessment is fundamental to preventative medicine. Without it, people may not be identified as being at high risk of a disease, and the chance to prevent that disease occurring may be lost. There are a variety of cardiovascular disease risk calculators available. Each has been developed from a different population and each has strengths and limitations. This article explores the important factors to consider when choosing which risk calculator to use, and outlines the pros and cons of calculators you may have access to in your practice.

Article points

1. Estimating cardiovascular disease (CVD) risk allows the identification of individuals without major risk factors who are actually at high risk. Identifying such people provides the opportunity to start treatments aimed at reducing CVD risk and means that therapies are provided fairly.
2. Risk calculators vary in their accuracy at predicting risk for an individual, but there appears to be reasonable agreement between them in ranking individuals according to risk.
3. Risk calculators help to identify individuals at high CVD risk allowing healthcare professionals to initiate risk reduction strategies.

Key words

- Cardiovascular disease
- Macrovascular disease
- Risk assessment
- Type 2 diabetes

Hermione Price is Clinical Research Fellow at the Diabetes Trials Unit, Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford.

Many people at increased risk of cardiovascular disease (CVD) are easy to identify because they have high blood pressure, high cholesterol, diabetes or smoke cigarettes. Estimating CVD risk allows the identification of individuals without major risk factors who are actually at high risk. Identifying such people provides the opportunity to start treatments aimed at reducing CVD risk and means that therapies are provided fairly.

What constitutes CVD?

Each calculator has a slightly different way of defining a CVD event, and hence CVD risk. The events that are usually included are myocardial infarction, stroke, and death from a myocardial infarction or stroke. Additionally, some calculators include angina, transient ischaemic attack, intermittent claudication, heart failure, coronary artery bypass grafting and percutaneous coronary angioplasty (Hippisley-Cox et al, 2007; D'Agostino et al, 2008). Essentially, the broader the definition of CVD, the more likely it is that an individual will have an event. This should be taken into account if two different calculators give very different estimates. It is also helpful to note that some calculators only estimate the risk of dying from a CVD event (Conroy et al, 2003).

Which calculator should I choose?

Are the people in the calculator like my patients?

Risk calculators are usually created from data collected from clinical trials or epidemiological studies. How well the individuals from the original study reflect your own patients will give you an idea as to whether this is the most appropriate calculator to use (Viljoen, 2008). For example, the very commonly used Framingham equations (Anderson et al, 1991; D'Agostino et al, 2008), recommended by NICE (2006) and found inside the back cover of the British National Formulary (Joint Formulary Committee, 2009) were developed using data collected from the Framingham studies.

Framingham is a town in Massachusetts in the US and has a mainly affluent White population. These equations may therefore be less applicable if your patients live in an area with high levels of deprivation or if you have

a high proportion of individuals from ethnic minority groups.

In addition, all studies have inclusion and exclusion criteria. It may not be appropriate to use a calculator to estimate risk for a 45-year-old if the study from which the calculator is derived only included people over the age of 50.

Is the calculator reliable?

To be sure that a risk calculator is applicable to most people it should have been shown to provide reliable or "externally validated" estimates for many different individuals. For example, its results should have been tested and validated in individuals from different ethnic groups, of different ages and of different socioeconomic status.

Many risk calculators only test their estimates by "internal validation", using a subset of the population from which the calculator was developed to verify the results. This does not ensure that the result will apply in a population that is different to the one from which it was developed.

Was enough information available to produce a reliable calculator?

All studies have inclusion and exclusion criteria for participants, and few studies will be able to collect every piece of information on every individual taking part. It is important, therefore, to know how much information was missing and what the creators of the calculator did about this.

Missing information will often be imputed (Hippisley-Cox et al, 2007). This is a statistical technique to replace missing data with "invented" data based on what information is available. If large quantities of data are imputed then this may make you feel uneasy about using a particular calculator.

What about people with diabetes?

Individuals with diabetes have a two- to four-fold greater risk of CVD than those without diabetes (Stamler et al, 1993). This increased risk is not reflected by all risk calculators (Coleman et al, 2007a). This can occur if only a small number of individuals with diabetes are included in the study from which the calculator is developed. For example, in the Framingham study only 428 of the 8491 individuals included in the calculator had diabetes.

Including more risk factors does not mean improved accuracy

Crude risk estimates can be made using simple measurements from the clinic, for example waist circumference or BMI (Dalton et al, 2003). Including additional measures, such as blood pressure and smoking status, and routine biochemical values, such as low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol, can improve the accuracy of these estimates (Anderson et al, 1991). However, many risk calculators attempt to increase the accuracy of their estimates further by including more and more novel risk factors (Woodward et al, 2007; Hippisley-Cox et al, 2008a). Interestingly, including additional risk factors rarely improves the accuracy of risk estimates any further. In addition, it can make many of these risk calculators redundant, as novel risk factors, including highly sensitive C-reactive protein levels or measures of carotid artery calcification, are not readily available in routine clinical practice (Ridker et al, 2002; Newman et al, 2008).

Although markers of obesity are often included in CVD risk equations in addition to routine biochemistry and blood pressure (Balkau et al, 2004; Cederholm et al, 2008), when combined with these risk factors, obesity no longer contributes to risk estimation. This is thought to be because the combined effects of other CVD risk factors outweighs the risk due to obesity (Coleman and Holman, 2007).

Population-based risk calculators

Framingham risk equations

Data from the landmark Framingham Heart Study and Framingham Offspring Study have been used to update the well-established Framingham risk equations (Anderson et al, 1991; D'Agostino et al, 2008). Framingham equations form the basis of many commonly used CVD risk prediction charts, including

Page points

1. To be sure that a risk calculator is applicable to most people it should have been shown to provide reliable or "externally validated" estimates for many different individuals.
2. If large quantities of data are imputed then this may make you feel uneasy about using a particular calculator.
3. Although markers of obesity are often included in cardiovascular disease risk equations in addition to routine biochemistry and blood pressure, when combined with these risk factors, obesity no longer contributes to risk estimation.
4. Framingham equations form the basis of many commonly used CVD risk prediction charts, including the New Zealand CVD risk tables. The equations are based on a population sample of 8491 adults aged 30–74 years.

Page points

1. Framingham only included a small number of individuals with diabetes (*n*=428), and the duration of known diabetes or other measures of disease severity such as HbA_{1c} are not included.
2. The Systematic Coronary Risk Evaluation (SCORE) risk equation was developed to estimate the 10-year risk of dying from a cardiovascular disease (CVD) event in the general population only.
3. Studies used in the SCORE risk equation did not have a standard method of recording diabetes so, unsurprisingly, these equations have been shown to underestimate risk in people with diabetes and should therefore not be used in people with diabetes.
4. The Diabetes Epidemiology: Collaborative Analysis of Diagnostic Criteria in Europe (DECODE) study group has developed risk equations to estimate the 5- and 10-year probability of dying from a CVD event. The equations were based on pooled data from 14 studies conducted across Europe.

the New Zealand CVD risk tables (Baker et al, 2000). The equations are based on a population sample of 8491 adults aged 30–74 years.

The advantage of Framingham over other population study-derived risk equations is that data were extremely well collected and CVD events were meticulously assessed.

Disadvantages include the dominance of the Caucasian population and the relatively high socioeconomic status of the participants. The Framingham investigators realised their population was not ethnically diverse and have now tested their equations in other ethnic groups, including African Americans, Native Americans, Japanese American men and Hispanic men (D'Agostino et al, 2001). The equations were found to give reasonable estimates of 5-year CVD risk in African American men without diabetes, but overestimated risk in all other ethnic groups. Despite these drawbacks, Framingham equations remain the CVD risk calculator of choice for NICE, and are generally accepted to provide reliable risk estimates in the general population (Stevens et al, 2005).

Framingham only included a small number of individuals with diabetes (*n*=428), and the duration of known diabetes or other measures of disease severity such as HbA_{1c} are not included. Both the original and more recent equations have been shown to underestimate CVD risk and coronary heart disease risk in individuals with diabetes, although the newer equations perform much better (Guzder et al, 2005; Price et al, 2008).

SCORE (Systematic Coronary Risk Evaluation)

The SCORE risk equation was developed to estimate the 10-year risk of dying from a CVD event in the general population only (Conroy et al, 2003). The equations were produced based on pooled data from 12 European cohort studies, including over 200 000 individuals and over 2.7 million patient years of follow-up.

Limitations of these equations include the fact that each risk factor was only measured once and not repeated over a period of time. SCORE also only includes a limited number of risk factors (age, sex, total cholesterol or total HDL–cholesterol ratio, smoking status and systolic blood pressure). While these are clearly the most important CVD risk factors, other key factors such as diabetes and ethnicity are absent.

The accuracy of the SCORE equations has been tested in a large Austrian cohort of over 44 000 people drawn from the general population (Ulmer et al, 2005). Of these, 487 died from a CVD event. SCORE predicted that 666 CVD deaths should have occurred. However, SCORE did correctly identify those most likely to die from a CVD cause.

Unfortunately, the studies used did not have a standard method of recording diabetes so, unsurprisingly, these equations have been shown to underestimate risk in people with diabetes (Coleman et al, 2007a), and should therefore not be used in people with diabetes.

DECODE (Diabetes Epidemiology: Collaborative Analysis of Diagnostic Criteria in Europe)

The DECODE study group has developed risk equations to estimate the 5- and 10-year probability of dying from a CVD event (Balkau et al, 2004). The equations were based on pooled data from 14 studies conducted across Europe. They include glucose categories based on fasting and 2-hour plasma glucose values, irrespective of whether or not an individual had a diagnosis of diabetes. The value of these equations in routine clinical practice, however, is questionable given the need for an oral glucose tolerance test.

The studies included were all good at determining if someone had died from a CVD event, and at the end of the study the investigators knew whether more than 95% of their participants were alive or dead. However, each study varied in how the biochemical information was collected.

The authors also did not know how many of their participants had a history of a CVD event before entering into the study. This makes it difficult to interpret the risk estimates because it is not clear if they are estimating the risk of having a first CVD event or the risk of

having another event. The investigators were also unable to include HDL cholesterol and triglycerides in their equations as these variables were not available. HDL has been shown to be an important CVD risk factor in other calculators (Coleman et al, 2007b, D'Agostino et al, 2008). Other limitations were that only half of the studies collected data for 10 years. This means that the 10-year risk estimates are based on much less information that the 5-year estimates. However, unlike many other studies, DECODE included a large number of younger people. Over half the male participants and one third of the female participants were under 50 years of age at entry to the study.

Two of the studies included in this calculator did not record if participants had diabetes, therefore it is not recommended for this population.

QRISK2

The QRISK2 investigators wanted to develop a risk equation reflecting ethnicity and social deprivation as these factors are not used in the Framingham equations.

QRISK2 (which supersedes QRISK) uses data collected from 531 UK general practices between 1993 and 2008 with data from 2.3 million people (Hippisley-Cox et al, 2008a). Participating practices were chosen because they used the Egton Medical Informations System.

QRISK2 holds data from more than 22 000 people of South Asian descent, 11 500 of Black African descent, 10 400 of Black Caribbean descent, and 19 700 of Chinese and other ethnic descent, and uses postcodes to determine economical status ("deprivation score"). QRISK2 also includes family history of premature CVD (a CVD event in a first-degree relative under 60 years of age), presence of rheumatoid arthritis, chronic kidney disease and atrial fibrillation.

An advantage of QRISK2 over QRISK is the link to the Office for National Statistics for collecting more accurate information on CVD deaths. QRISK2 has been found to be better than both QRISK and the Framingham equations in predicting the chance of having a CVD event (Hippisley-Cox et al, 2008b). However, although QRISK2 uses information from a large sample of the general population, it does have several limitations. It excluded all individuals taking statin therapy, those with no valid deprivation score, temporary residents and those without at least 1 year of follow-up.

The exclusion of those already taking statin therapy is interesting as they would have already been identified as being at high CVD risk. The exclusion of those without a valid deprivation score or temporary residents may exclude some of the most deprived communities, including those who do not have a fixed address or travelling communities.

A particular problem with QRISK was the large amount of missing data, particularly for total cholesterol (Hippisley-Cox et al, 2007). Missing data remained an issue for QRISK2 but investigators used a more sophisticated statistical technique to replace missing data. QRISK2 has been internally validated but has not been tested in any other populations, whereas the original QRISK was successfully externally validated using data from The Health Improvement Network (Hippisley-Cox et al, 2008b).

Interestingly, although QRISK2 was developed specifically to improve CVD risk estimation in deprived and ethnic minority groups, self-reported ethnicity was only available for a quarter of participants in the study.

Type 2 diabetes is included in QRISK2 as "yes/no" and is therefore likely to underestimate risk in people with diabetes, as other calculators that do not include HbA_{1c}, or duration of known diabetes, tend to underestimate risk in people with diabetes (Coleman et al, 2007a; Price et al, 2009).

ASSIGN (Assessing CVD Risk Using Scottish Intercolliegiate Guideline Network)

The aim of the ASSIGN equations was to reduce health inequalities by including a measure of social deprivation in CVD risk calculations (Woodward et al, 2007). The ASSIGN score investigators felt that the Framingham equations do not adequately address the gradient in CVD risk associated with increasing social deprivation and that this may result in an inequitable allocation of resources.

Page points

1. The QRISK2 investigators wanted to develop a risk equation reflecting ethnicity and social deprivation as these factors are not used in the Framingham equations.
2. QRISK2 (which supersedes QRISK) uses data collected from 531 UK general practices between 1993 and 2008 with data from 2.3 million people.
3. Although QRISK2 uses information from a large sample of the general population it does have several limitations. It excluded all individuals taking statin therapy, those with no valid deprivation score, temporary residents and those without at least 1 year of follow-up.
4. The aim of the ASSIGN (Assessing CVD Risk Using Scottish Intercolliegiate Guideline Network) equations was to reduce health inequalities by including a measure of social deprivation in cardiovascular disease risk calculations.

Page points

1. The Swedish National Diabetes Register calculator is diabetes-specific and based on data from over 11 000 individuals in the Swedish National Diabetes Register (Cederholm et al, 2008).
2. The UKPDS (United Kingdom Prospective Diabetes Study) Risk Engine was developed specifically for estimating cardiovascular disease (CVD) risk in individuals with type 2 diabetes using data from the UKPDS clinical trial.
3. Version three of the UKPDS Risk Engine calculates the probability of having a CVD event and the probability of dying from a CVD event in the next 10 years.
4. For the general population (without diabetes) the Framingham-derived CVD equations provide accurate estimates of CVD risk. In individuals with diabetes a risk calculator designed specifically for use in individuals with diabetes should be used, for example the UKPDS Risk Engine.

ASSIGN uses data from the Scottish Heart Health Extended Cohort, which included over 13 000 adults who had not had a CVD event when they entered the study (Woodward et al, 2007). Participants were followed for between 10 and 21 years and during this time there were over 1000 CVD events. These equations have not been tested in any other population but ASSIGN was found to be in good agreement with, but no better than, Framingham (Woodward et al, 2007). The ASSIGN equations were unable to include ethnicity as their study did not include many people from different ethnic groups.

Until these equations have been tested and shown to work in people with diabetes they should not be used.

Risk calculators specifically designed for people with diabetes

Swedish National Diabetes Register calculator

The Swedish National Diabetes Register calculator is diabetes specific and based on data from over 11 000 individuals in the Swedish National Diabetes Register (Cederholm et al, 2008).

This calculator includes over 1400 CVD events, although it was developed using routinely collected data and not via a clinical trial. This means that information has not been collected consistently and that the CVD events were not independently verified. We cannot therefore be sure that the data used is accurate. Additionally, the calculator has not been externally validated.

UKPDS (United Kingdom Prospective Diabetes Study) Risk Engine

The UKPDS Risk Engine was developed specifically for estimating CVD risk in individuals with type 2 diabetes using data from the UKPDS clinical trial. Version three of the Risk Engine (Coleman et al, 2007b) calculates the probability of having a CVD event and the probability of dying from a CVD event in the next 10 years.

The UKPDS Risk Engine includes data from 3475 individuals enrolled in the UKPDS study. Participants had newly diagnosed type 2 diabetes and had not had a CVD event when they entered the study (Coleman et al, 2007b). The Risk Engine includes HbA_{1c} levels, and takes into account duration of known diabetes. It has been tested against data collected from the CARDS (Collaborative Atorvastatin Diabetes Study) (Colhoun et al, 2004), but has also been found to underestimate risk in individuals with diabetes from the UK, although not as much as the Framingham risk equations (Guzder et al, 2005). This underestimation may have occurred because this study did not include individuals who had subclinical (silent) coronary events.

The Risk Engine has undergone external validation and has been shown to be more appropriate for use in individuals with diabetes than calculators developed from samples of the general population. Interestingly, the Engine has also been found to be capable of ranking CVD risk even in individuals without diabetes. However, it overestimated risk in individuals without diabetes when compared with the Framingham equations (Simmons et al, 2009).

Conclusion

Risk calculators help to identify individuals at high CVD risk, allowing healthcare professionals to initiate risk reduction strategies. Risk calculators vary in their accuracy at predicting risk for an individual, but there appears to be reasonable agreement between them in ranking individuals according to risk. This information is clearly useful in determining provision of resources but is less useful to the individual patient who wants an accurate estimate of their own personal level of risk. Applying the wrong risk calculator, particularly in those with diabetes, could result in an individual being inappropriately denied risk-reducing treatment (Price et al, 2009).

For the general population (without diabetes) the Framingham-derived CVD equations provide accurate estimates of CVD risk. In individuals with diabetes a risk calculator designed specifically for use in those with the condition should be used, for example the UKPDS Risk Engine. ■

Anderson K, Odell P, Wilson P, Kannel W (1991) Cardiovascular disease risk profiles. *Am Heart J* **121**: 293–8

Baker S, Priest P, Jackson R (2000) Using thresholds based on risk of cardiovascular disease to target treatment for hypertension: modelling events averted and number treated. *BMJ* **320**: 680–85

Balkau B, Hu G, Qiao Q et al (2004) Prediction of the risk of cardiovascular mortality using a score that includes glucose as a risk factor. The DECODE Study. *Diabetologia* **47**: 2118–28

Cederholm J, Eeg-Olofsson K, Eliasson B (2008) Risk prediction of cardiovascular disease in type 2 diabetes: a risk equation from the Swedish National Diabetes Register. *Diabetes Care* **31**: 2038–43

Coleman R, Holman R (2007) Are measures of adiposity associated independently with incident cardiovascular disease in individuals with newly diagnosed type 2 diabetes? *Diabetologia* **50**(Suppl 1): S502

Coleman R, Stevens R, Retnakaran R, Holman R (2007a) Framingham, SCORE, and DECODE risk equations do not provide reliable cardiovascular risk estimates in type 2 diabetes. *Diabetes Care* **30**: 1292–3

Coleman R, Stevens R, Holman R (2007b) Estimating cardiovascular disease risk for individuals with type 2 diabetes: the new UKPDS Risk Engine. *Diabet Med* **24**(Suppl 1): 56

Colhoun H, Betteridge D, Durrington P et al (2004) Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. *Lancet* **364**: 685–96

Conroy RM, Pyörälä K, Fitzgerald AP et al (2003) Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* **24**: 987–1003

Dalton M, Cameron A, Zimmet P et al (2003) Waist circumference, waist-hip ratio and body mass index and their correlation with cardiovascular disease risk factors in Australian adults. *J Intern Med* **254**: 555–63

D'Agostino R, Grundy S, Sullivan L et al (2001) Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA* **286**: 180–7

D'Agostino R, Vasan R, Pencina M et al (2008) General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* **117**: 743–53

Guzder R, Gatling W, Mullee M et al (2005) Prognostic value of the Framingham cardiovascular risk equation and the UKPDS risk engine for coronary heart disease in newly diagnosed Type 2 diabetes: results from a United Kingdom study. *Diabet Med* **22**: 554–62

Hippisley-Cox J, Coupland C, Vinogradova Y et al (2007) Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* **335**: 136

Hippisley-Cox J, Coupland C, Vinogradova Y et al (2008a) Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* **336**: 1475–82

Hippisley-Cox J, Coupland C, Vinogradova Y (2008b) Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. *Heart* **94**: 34–9

Joint Formulary Committee (2009) *British National Formulary 57*. BMJ Group and RPS Publishing, London

Newman A, Naydeck B, Ives D et al (2008) Coronary artery calcium, carotid artery wall thickness, and cardiovascular disease outcomes in adults 70 to 99 years old. *Am J Cardiol* **101**: 186–92

NICE (2006) *Statins for the Prevention of Cardiovascular Events in Patients at Increased Risk of Developing Cardiovascular Disease or those with Established Cardiovascular Disease*. NICE, London

Price HC, Tucker L, Griffin SJ, Holman RR (2008) The impact of individualised cardiovascular disease (CVD) risk estimates and lifestyle advice on physical activity in individuals at high risk of CVD: a pilot 2x2 factorial understanding risk trial. *Cardiovasc Diabetol* **7**: 21

Price H, Coleman R, Stevens R, Holman R (2009) Impact of using a non-diabetes-specific risk calculator on eligibility for statin therapy in type 2 diabetes. *Diabetologia* **52**: 394–97

Ridker P, Rifai N, Rose L et al (2002) Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. *N Engl J Med* **347**: 1557–65

Simmons R, Price H, Holman R et al (2009) Performance of the UK Prospective Diabetes Study Risk Engine and the Framingham Risk Equations in estimating cardiovascular disease in the EPIC-Norfolk Cohort. *Diabetes Care* **32**: 708–13

Stamler J, Vaccaro O, Neaton J, Wentworth D (1993) Diabetes, other risk factors, and 12-yr cardiovascular mortality for men screened in the Multiple Risk Factor Intervention trial. *Diabetes Care* **16**: 434–44

Stevens R, Coleman R, Holman R (2005) Framingham risk equations underestimate coronary heart disease risk in diabetes. *Diabet Med* **22**: 228

Ulmer H, Kollerits B, Kelleher C (2005) Predictive accuracy of the SCORE risk function for cardiovascular disease in clinical practice: a prospective evaluation of 44 649 Austrian men and women. *Eur J Cardiovasc Prev Rehabil* **12**: 433–41

Viljoen A (2008) Cardiovascular risk estimation - making sense of the numbers. *Int J Clin Pract* **62**: 1300–3

Woodward M, Brindle P, Tunstall-Pedoe H (2007) Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart* **93**: 172–6

"Applying the wrong risk calculator, particularly in individuals with diabetes, could result in an individual being inappropriately denied risk-reducing treatment."