When commenting please quote page number and line number for each comment

Section 4.3 of the Guideline on Cardiovascular Risk Assessment: the modification of blood lipids for the primary and secondary prevention of cardiovascular disease

This guideline has already been the subject of public consultation. This document contains a revised version of section 4.3, which is being issued for a second consultation. This is because a new equation, QRISK, for the estimation of cardiovascular risk has been developed. Emerging evidence suggests that QRISK gives a better estimation of risk in the general population of England and Wales than the Framingham equations. The GDG has reviewed this evidence and has revised its recommendations on cardiovascular risk assessment in this document. NICE commissioned expert reviews of QRISK and these are given in the appendix.

Consultation Version

February 2008

National Collaborating Centre for Primary Care

Contents

4 Identification and assessment of people at high risk of cardiovascular disease (CVD)

4.3 Assessment of cardiovascular risk

4.3.1 Introduction

Estimates of CVD risk derived from equations are not an exact science but are better than clinical judgment alone for the estimation of CVD risk.

A number of risk assessment equations are available that estimate cardiovascular risk in individuals. They have been derived from studies of individuals who have been followed up often for substantial lengths of time. Risk assessment equations predict risk best in the type of population from which they were derived. Equations derived from North American populations from the 1960s to the 1980s when coronary heart disease (CHD) was at its peak overestimate risk in contemporary European populations by around twofold in Southern European populations and by 50% or more in Northern European populations including the UK. Conversely, such equations may underestimate risk in populations such as people with diabetes, South Asian men or the most socially deprived who are at higher than average risk.

4.3.2 Recommendations for assessment of cardiovascular risk

4.3.2.1 CVD risk should be calculated using the published QRISK equation.

4.3.2.2 The QRISK risk equation should not be used for people known to have:

- *coronary heart disease/angina*
- *stroke/transient ischaemic attack*
- *peripheral arterial disease.*

4.3.2.3 The QRISK risk equation should not be used to reassess risk in people previously identified as at high risk of CVD and who are already on treatment.

4.3.2.4 The QRISK risk equation should not be used for people who because of their condition are already considered at high risk of CVD because of:

- *familial hypercholesterolaemia or other monogenic disorders of lipid metabolism*
- *diabetes. (see the forthcoming NICE clinical guideline 'Type 2 diabetes: the management of type 2 diabetes (update)'. Publication expected April 2008. Information available from www.nice.org.uk)*

4.3.2.5 If the risk estimate is marginally below the threshold, clinical judgement should be used to determine whether further treatment of risk factors should be offered (for example, South Asian males)

4.3.2.6 Cardiovascular risk scores may not be appropriate in people who are at increased CVD risk due to underlying medical conditions or treatments. These include people treated for HIV or with anti-psychotic medication, people with chronic kidney disease and patients with autoimmune disorders such as SLE and rheumatoid arthritis.

4.3.2.7 People aged 75 years and over should be considered to be at increased risk of CVD, particularly people who smoke or who have raised blood pressure, and they are likely to benefit from statin treatment. Assessment and treatment should be guided by the benefits and risks of treatment for the individual, informed preference of the person and co-morbidities that may make such treatment inappropriate

4.3.2.8 People in whom familial hypercholesterolaemia or other familial disorders are suspected because of a combination of clinical

findings, lipid profiles and family history of premature CHD (see the forthcoming NICE clinical guideline 'Familial hypercholesterolaemia: the identification and management of adults and children with familial hypercholesterolaemia'. Publication expected August 2008. Information available from www.nice.org.uk)

4.3.2.9 People with severe hyperlipidaemia should be considered for further investigation and/or specialist review.

4.3.3 Evidence statements for assessment of cardiovascular risk

4.3.3.1	Different risk assessment methods exist. The most widely used and researched are derived from the Framingham cohort.
4.3.3.2	In representative populations, recognised Framingham-based methods offer reasonable discrimination between high- and low-risk individuals but tend to overestimate the absolute risk of CVD in lower risk populations and underestimate risk in high-risk populations. There has been concern that estimates derived from North American populations dating back 30 years may not accurately estimate risk in contemporary European populations when CHD mortality has fallen by more than half during this period. Overall the Framingham risk equation is likely to overestimate risk in the current UK population, more so in Southern England than Northern England or Scotland.
4.3.3.3	Framingham-based methods may underestimate risk in people at high risk such as people with a strong family history of premature CVD, certain ethnic groups and those from relatively socio-economically deprived backgrounds. They may also underestimate risk in people with extreme risk factors or other clinical risks not included in the model.
4.3.3.4	There are no consistent differences in the generalisability of one Framingham model over another.
4.3.3.5	The following endpoints have been used by the statin technology appraisal report to establish treatment thresholds: fatal and non-fatal myocardial infarction, acute coronary syndrome, stable angina, stroke, and transient ischaemic attacks. (NICE technology appraisal 94, 'Statins for the prevention of cardiovascular events').
4.3.3.6	When used in conjunction with the Framingham estimates, those defined by the NICE Technology Appraisal are the most

	appropriate. When considering management strategies based on other risk equations, endpoints such as revascularisation, peripheral arterial disease and other disease processes associated with atherosclerosis may also be relevant.
4.3.3.7	Framingham based risk scoring methods do not accurately estimate risks in some groups of people.
4.3.3.8	Several risk factors have not been included in the Framingham risk equations and some adjustment of this risk estimate may be required to more accurately represent an individual's absolute risk: • Family history of a premature event from CVD: first-degree male relatives under the age of 55 years and first-degree female relatives under the age of 65 years • Ethnic group • Socio-economic status • People already on treatment that modifies CV risk • Extremes of risk factors, for example people who have a body mass index over 40 kg/m^2.
4.3.3.9	There are differences in cardiovascular risk between black and minority ethnic groups and the white population in England and Wales.
4.3.3.10	For men, the risk of CVD was higher in South Asian ethnic groups (with some subgroup heterogeneity) than for men in the white population.
4.3.3.11	For men there is no robust evidence for a difference in the risks of CVD other than that between South Asian ethnic groups and the general population.
4.3.3.12	For women there is no robust evidence for a difference in the risks of CVD between South Asian ethnic groups (with considerable

	subgroup heterogeneity) and the general population.
4.3.3.13	There is increased risk of CVD in people with a family history of premature CVD.
4.3.3.14	Cohort studies have shown a consistent association between having a positive family history of CVD and an increased risk of developing CVD. This risk remains even when adjusted for age, social class, body mass index, systolic blood pressure, blood lipids (cholesterol, triglycerides), fasting glucose and smoking status. The exact relative risk varies according to sex and nature of relationship between the individual with premature CVD and the index case.
4.3.3.15	The younger the age at which the family event occurred and the greater the number of family members involved, the greater the relative risk.
4.3.3.16	Cardiovascular risk is closely associated with socio-economic status. Framingham equations do not include socio-economic status and underestimate risk in people who are relatively socially deprived. The use of equations that do not include a measure of socio-economic status may exacerbate inequalities in CVD.
4.3.3.17	QRISK is a new risk score that has been developed using routine data from UK electronic primary care patient records.
4.3.3.18	QRISK includes social deprivation, family history, body mass index and antihypertensive treatment that are not included in the Framingham equation.
4.3.3.19	QRISK has better discrimination in a UK population than Framingham
4.3.3.20	QRISK is better calibrated to the UK population than Framingham

4.3.3.21	Little evidence was found supporting or refuting the assumption that cardiovascular risk assessment by clinicians improves health outcomes. The interventions showed no improvement in predicted absolute cardiovascular risk or in declared primary outcomes.
4.3.3.22	A study in hypertensive patients has shown a small reduction in systolic blood pressure associated with the use of a risk chart but not when used in conjunction with a computer based clinical decision support system.
4.3.3.23	Another study has shown very low uptake of risk-scoring methods by clinicians that would have obscured any beneficial effect on blood pressure by the intervention.
4.3.3.24	The accuracy of use of chart based systems has been questioned. Current evidence is an insufficient basis on which to judge the effectiveness of CVD risk estimation as a method of improving health outcomes.

4.3.4 Methods for multiple risk factor assessment to estimate absolute cardiovascular risk in people who are at risk of CVD

A recent systematic review (Beswick, A. D., Brindle, P., Fahey, T. et al) (Appendix J of the full guideline) was used as the evidence source. Literature searching beyond the search date of the systematic review identified two further risk scores developed in UK populations (QRISK discussed in section 4.3.5, and ASSIGN discussed in sections 4.3.4.5 and 4.3.5). The Beswick *et al* systematic review compared the accuracy of risk scoring methods such as charts and tables compared with full prediction models, namely, the Framingham-Anderson model of 1991(Anderson, K. M., 1991). A complete reference to the materials and evidence reviewed is given in Appendix J of the full guideline.

Eleven derived risk charts, tables and nomograms were identified comparing risk calculations with the original Framingham-Anderson prediction model (1991).

The tools identified were as follows:

- Sheffield tables (2 versions) (Haq, I. U., Jackson, P. R., Yeo, W. W. et al , 1995) (Ramsay, L. E., Haq, I. U., Jackson, P. R. et al , 1996) (Wallis, E. J., Ramsay, L. E., Ul, HaqI et al , 2000)

- Joint British Societies (JBS) charts (2 versions) (Joint British recommendations on prevention of coronary heart disease in clinical practice: summary. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, British Diabetic Association, 2000) (Joint British recommendations on prevention of coronary heart disease in clinical practice. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, endorsed by the British Diabetic Association, 1998)

- Joint European Societies (JBS) charts (2 versions) (Wood, D., De, Backer G, Faergeman, O. et al , 1998) (Conroy, R. M., Pyorala, K., Fitzgerald, A. P. et al , 2003)

- Canadian nomograms (McCormack, J. P., Levine, M., and Rangno, R. E., 1997)

- New Zealand charts (3 versions) (1996 National Heart Foundation clinical guidelines for the assessment and management of dyslipidaemia. Dyslipidaemia Advisory Group on behalf of the scientific committee of the National Heart Foundation of New Zealand., 1996) (McLeod, A. J. and Armitage, M., 1998) (Jackson, R., 2000)

- World Health Organization and the International Society for Hypertension (WHO-ISH) chart http://www.ish-world.com/default.aspx?Guidelines.

It was found that the early versions of the Sheffield Tables (Haq, I. U., Jackson, P. R., Yeo, W. W. et al , 1995) (Ramsay, L. E., Haq, I. U., Jackson,

P. R. et al , 1996) and the Joint European Societies charts (Wood, D., De, BackerG, Faergeman, O. et al , 1998) (Conroy, R. M., Pyorala, K., Fitzgerald, A. P. et al , 2003) had poor sensitivity as they did not include individual values for HDL cholesterol in the risk calculation. More recent Sheffield tables (Wallis, E. J., Ramsay, L. E., Ul, HaqI et al , 2000) and Joint British Society charts (Joint British recommendations on prevention of coronary heart disease in clinical practice: summary. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, British Diabetic Association, 2000) (Joint British recommendations on prevention of coronary heart disease in clinical practice. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, endorsed by the British Diabetic Association, 1998) show reasonable sensitivity and specificity compared with the full Framingham Anderson model. The 1997 Canadian nomograms (McCormack, J. P., Levine, M., and Rangno, R. E., 1997) included HDL cholesterol in their risk calculation however they were very poor at identifying patients at high levels of risk. The WHO-ISH 1999 table suffers from generalisation of the Framingham-Anderson model with risk factor counting substituting for continuous clinical variables. The New Zealand charts have only moderate sensitivity and specificity and provide assessment of CVD risk (1996 National Heart Foundation clinical guidelines for the assessment and management of dyslipidaemia. Dyslipidaemia Advisory Group on behalf of the scientific committee of the National Heart Foundation of New Zealand., 1996) (McLeod, A. J. and Armitage, M., 1998) (Jackson, R., 2000). The most recent Joint British Society charts estimate CVD risk but were not available at the time of this review.

In conclusion, the systematic review by Beswick *et al* (Beswick, A. D., Brindle, P., Fahey, T. et al) (Appendix J of the full guideline) showed that comprehensive information is required in risk tables and charts. The inclusion of HDL cholesterol gives the most accurate estimate of cardiovascular risk.

4.3.4.1 Endpoints used for assessment when estimating cardiovascular risk

The choice of CVD endpoint is important as it affects the numbers of people reaching treatment thresholds and the numbers targeted for risk reduction treatments.

The endpoints recommended in this guideline are the same as those used in the NICE Technology Appraisal 94: Statins for the prevention of cardiovascular events (2006). The scope for this guideline includes risk factor modification for symptomatic atherosclerotic vascular disease including revascularisation and peripheral arterial disease and these endpoints should be included where appropriate in other recommended risk equations.

Adjusting Framingham risk equations – these sections will be in an appendix in the final guideline.

4.3.4.2 Adjusting the calculated Framingham cardiovascular risk estimate by other risk factors

A systematic review by Brindle *et al* (Brindle, P. M., Beswick, A. D., Fahey, T. et al , 2006) reviewed the accuracy of Framingham-based methods to estimate risk in populations other than those in which the models were derived (external validation).

Data were extracted on the ratio of the predicted to the observed 10-year risk of CVD and CHD from 27 studies with data from 71,727 participants. These studies used either the Framingham-Anderson (1991) (Anderson, K. M., 1991) or Wilson (Wilson, P. W. F., D'Agostino, R. B., Levy, D. et al , 1998) risk scores (methods using the outcomes of combined fatal and non-fatal CHD or CVD) and covered a wide range of different population groups: Populations varied in nationality, age range and sex, date of recruitment and outcomes studied. The groups studied were representative samples of men and women, people with diabetes, people with raised cholesterol, people on treatment for hypertension, patients with no CHD determined by angiography and patients with a family history of CVD.

For CHD, the predicted to observed ratios ranged from 0.43 in a study of people with a family history of CHD (that is, predicting a lower risk than was observed) to 2.87 in a study of women from Germany (PROCAM) (that is, predicting a much higher risk than was observed) (Hense, H. W., Schulte, H., Lowel, H. et al , 2003). Under-prediction was observed in studies of higher risk patients such as those with diabetes, a strong family history of premature CVD, people from geographical areas with a high incidence of disease and people in socio-economically deprived groups.

For CVD, there was similar trend of increasing under-prediction with increasing risk of the population.

Over-prediction of risk occurs when Framingham equations are applied to populations with a lower baseline risk than that experienced by the Framingham cohort. Over-prediction was seen in lower and medium risk primary care and occupational populations in Germany (Hense, H. W., Schulte, H., Lowel, H. et al , 2003), France and Northern Ireland (Empana, J. P., Ducimetiere, P., Arveiler, D. et al , 2003) and a US screening cohort with a medium level of observed risk (Greenland, P., La Bree, L., Azen, S. P. et al , 2004). In the multicentre clinical trial of Bastuji-Garin *et al*, CHD risk was over-estimated and this was seen across eight Western European countries and Israel (Bastuji-Garin, S., Deverly, A., Moyse, D. et al , 2002). Within England, Wales and Scotland, over-prediction by the Framingham equations occurred in all regions but was greater in the South and the Midlands/Wales where there was relatively lower mortality and morbidity than in Scotland and the North of England (Brindle, P., Emberson, J., Lampe, F. et al , 2003).

This systematic review shows that the accuracy of the Framingham risk estimates cannot be assumed, and that it relates to the background risk of CVD in the population to which it is being applied. Over-estimation of risk tends to occur in populations with low observed risk and underestimation in high-risk groups.

4.3.4.3 Adjustment of the Framingham cardiovascular risk score to take account of ethnicity

The rates of CVD vary between ethnic groups; however, the Framingham risk score does not take ethnicity into account as a risk factor.

Studies were identified which provide evidence for differences in risk by ethnic group in the UK and the need to adjust risk estimates to take into account ethnic origin when estimating an individual's risk of CVD (Cappuccio, F. P., Oakeshott, P., Strazzullo, P. et al , 2002) (Quirke, T. P., Gill, P. S., Mant, J. W. et al , 2003) .

The method of adjustment was considered in three papers. Bhopal et al's (Bhopal, R., Fischbacher, C., Vartiainen, E. et al , 2005) paper included 6448 men and women aged 25 to 74 years from the Newcastle Health and Lifestyle Survey. The hazard ratio adjusted for age and sex for CHD death in South Asians combined compared with Europeans was 2.23 (95% CI 1.13 to 4.38), the corresponding ratio for stroke mortality was 1.35 (95% CI 0.32 to 5.7).

A study by Aarabi and Jackson (Aarabi, M. and Jackson, P. R., 2005) used risk factor data from 4497 individuals identified from the Health Surveys for England 1998 and 1999, who were eligible to have their risk of a first CHD event calculated by the Framingham equation. Arabi and Jackson considered adding 10 years to the age of South Asian people as the simplest way of calculating CHD risk using paper based methods. The validity of this method, which assumes an excess risk of 1.79, is uncertain.

The study by Brindle et al (Brindle, P., May, M., Gill, P. et al , 2006) included 3,778 men and 4544 women aged 35 to 54 years from the Health Surveys for England 1998 and 1999 and the Wandsworth Heart and Stroke Study, both of which are community-based surveys. The authors estimated the incidence rate from prevalence data for 7 minority ethnic groups: Indians, Pakistanis, Bangladeshis, black Caribbean, Chinese (from the Health Surveys for England 1998/99) and black Africans (from the Wandsworth Heart and Stroke Study). The incidence rate was estimated because of the lack of prospective data on British black and minority ethnic groups.

The sex-specific and age-standardised prevalence ratio for CHD and for CVD for each ethnic group compared with the general British population was obtained from the Health Surveys for England 1998/99. Separate risk estimates were developed for CHD and CVD for both men and women for each ethnic group.

Calculated age-adjusted CVD prevalence ratios for seven ethnic groups showed considerable variation. In men, the highest ratio was observed in Bangladeshis (HR1.39, CI 0.82 to1.96) and the lowest among Chinese (HR0.49, CI 0.16 to 0.82); in women, the highest ratio (HR1.33, CI 0.70 to 1.96) was in Pakistanis and the lowest (HR0.22, CI 0 to 0.53) among Chinese.

This model has not been validated.

In summary, there is consistent evidence to support the need for adjustment of Framingham risk estimates to take account of ethnicity in UK populations but the best method for achieving this remains uncertain. Current guidance by the Joint British Societies (JBS2) (Wood, D., Wray, R., Poulter, N. et al , 2005) recommends multiplying the Framingham score by a correction factor of 1.4 for South Asian people; however, this does not acknowledge the difference between the sexes. There are particular problems in estimating risk for people of Afro-Caribbean origin who have a higher risk of stroke but a lower risk of ischemic heart disease.

It was noted that the determination of ethnicity itself is problematic despite much debate (Gill, P. S., Kai, J., Bhopal, R. S. et al , 2007). It is a multidimensional concept and embodies one or more of the following: ‘shared origins or social background; shared culture and traditions that are distinctive, maintained between generations, and lead to a sense of identity and group; and a common language or religious tradition’. For pragmatic reasons the self-determined Census question on ethnic group is acceptable. South Asian is a broad category and is generally defined as people assigning themselves as Indian, Pakistani, Bangladeshi and Sri Lankans.

The GDG agreed with the data compiled by Brindle *et al* (Brindle, P., May, M., Gill, P. et al , 2006) that indicated that a risk estimate 1.4 times that of the white population was the most appropriate weighting to use for adjustment of the Framingham equation in men of South Asian origin. There was no significant increase in risk among South Asian women. Although some other ethnic groups had low levels of risk in comparison to white people, this was not sufficiently robust on which to base a recommendation.

4.3.4.4 Adjustment of the Framingham cardiovascular risk score to take into account family history

Three studies were found addressing the extent to which family history predicts risk. These studies are the Framingham Offspring Study by Lloyd-Jones et al (Lloyd-Jones, D. M., Nam, B. H., D'Agostino, R. B., Sr. et al , 2004) the Malmo Preventive Project (MPP) by Nilsson et al (Nilsson, P. M., Nilsson, J. A., and Berglund, G., 2004) (follow up study) and the Physicians' Health Study (PHS) and the Women's Health Study (WHS) (Sesso, H. D., Lee, I. M., Gaziano, J. M. et al , 2001).

The Framingham Offspring Study

Lloyd-Jones *et al* (Lloyd-Jones, D. M., Nam, B. H., D'Agostino, R. B., Sr. et al, 2004) determined whether parental CVD predicts offspring events independent of traditional risk factors. The population consisted of 2302 men and women with a mean age of 44 years in the Framingham Offspring Study, who were free of CVD and whose parents were both in the original Framingham cohort. The authors examined the association of parental CVD with an 8-year risk of offspring CVD using pooled logistic regression.

Compared with the participants with no parental CVD, those with at least 1 parent with premature CVD (onset age < 55 years in father, < 65 years in mother) had a greater risk for events, with age-adjusted odds ratios of 2.6 (95% CI 1.7 to 4.1) for men and 2.3 (95% CI 1.3 to 4.3) for women. Multivariate adjustment resulted in odds ratios of 2.0 (95% CI 1.2 to 3.1) for men and 1.7 (95% CI 0.9 to 3.1) for women. Non-premature parental CVD and parental coronary disease were weaker predictors.

The Malmo Preventive Project (MPP)

Nilsson et al (Nilsson, P. M., Nilsson, J. A., and Berglund, G., 2004) studied the adjusted relative risk of CVD events in offspring of parents with cardiovascular mortality before 75 years. A total of 22 444 men and 10 902 women attended a screening programme between 1974 and 1992 and were followed up through national record linkage. There was an increased risk of CVD events (mortality and morbidity) in offspring in relation to a positive family history of parental CVD mortality before 75 years. The multivariate adjusted relative risk (RR) for father-son heritage was 1.22 (95% CI 1.02 to 1.47; $P < 0.05$), for mother-son heritage, RR = 1.51 (95% CI 1.23 to 1.84, $P < 0.001$), for father-daughter heritage, RR = 1.20 (95% CI 0.83 to 1.73) and for mother-daughter heritage, RR = 0.87 (95% CI 0.54 to 1.41).

Subdividing parental age of early death into age groups 50-68, 69-72 and 73-75 years showed a graded association for maternal influence: RR = 1.82 (95% CI 1.35 to 1.46), 1.55 (95% CI 1.14 to 2.10) and 1.50 (95% CI 1.13 to 1.98) respectively but not for paternal influence, RR 1.29 (95% CI 0.99 to 1.69), 1.08 (95% CI 0.81 to 1.44) and 1.40 (95% CI 1.12 to 1.76) respectively using surviving parents or mortality after 75 years as the reference group.

The Physicians' Health Study (PHS) and the Women's Health Study (WHS)

Sesso *et al* (Sesso, H. D., Lee, I. M., Gaziano, J. M. et al , 2001) prospectively studied 22 071 men from the Physicians' Health Study (PHS) and 39 876 women from the Women's Health Study (WHS) with data on parental history and age at MI.

Compared with men with no parental history, those with a maternal, paternal and both maternal and paternal history of MI had a RR of CVD of 1.71, 1.40 and 1.85 respectively; among women, the corresponding RRs were 1.46, 1.15 and 2.05 respectively.

Sesso *et al* (Sesso, H. D., Lee, I. M., Gaziano, J. M. et al , 2001) also looked at the effect of parental age: For men, maternal age at MI of < 50, 50 to 59, 60

to 69, 70 to 79 and ≥ 80 years had RRs of 1.00, 1.88, 1.88, 1.67 and 1.17. For women, the RRs for maternal age at MI of < 50, 50 to 59 and ≥ 60 years were 2.57, 1.33 and 1.52. Paternal age at MI of < 50, 50 to 59, 60 to 69, 70 to 79 and ≥ 80 years in men had RRs of 2.19, 1.64, 1.42 1.16 and 0.92; in women, for paternal age at MI of < 50, 50 to 59 and ≥ 60 years, the RRs were 1.63, 1.33 and 1.13.

The GDG noted that there was a continuous distribution of risk, which tended to increase the younger the age at which the family member had an event. Increased risk was noted to be present even up to age 75 years. The number of family members was also related to risk, and risk was greater where female relatives were affected. For simplicity the GDG considered that risk should be adjusted by 1.5 where there was a female first-degree relative under 65 years with CHD or a first-degree male relative under 55 years. Additional family members in this category would further increase risk. If more than one first-degree relative is affected, the risk estimate should be increased by a factor of up to 2.0.

4.3.4.5 Adjustment of the Framingham cardiovascular risk score to take into account socio-economic status

There is a widening relative gap in mortality and morbidity associated with socio-economic status. There has been a substantial reduction in CVD in the past two decades but the poorer sections of society have not improved as fast as the more affluent. In 1986 to 1992 mortality from circulatory disease was 69% greater in people from social classes IV and V than that in people in social classes I and II and by 1997 to 1999 this had increased to 86% (White, C., von Galen, F., and Chow, Y. H., 2003). This represents a decrease between socio-economic groups in absolute mortality difference but a widening of the relative difference. This relative inequality has been a cause for governmental concern and tackling health inequalities in CVD is a major component of current governmental strategy (Department of Health, 2003). Mortality from circulatory diseases in the most deprived category is currently threefold higher in women and 2.7 times higher in men than in the least deprived category.

During the course of this guideline development the Scottish ASSIGN score has been published and adopted as part of SIGN guidance but at the time of writing had not been validated in an English or UK population. It was developed in a Scottish cohort. In this cohort the Framingham score overestimated risk overall and in each quintile of social deprivation. It substantially underestimated the variation in risk with deprivation. The relative risk of observed 10-year CVD risk (sexes combined) analysed across population fifths had a steep gradient, from least to most deprived, of 1.00, 1.81, 1.98, 2.22, and 2.57. Expected risk, calculated from baseline risk factor values and the Framingham score, had one quarter of that gradient, with relative risks of 1.00, 1.17, 1.19, 1.28, and 1.36 (Woodward, M., Brindle, P., Tunstall-Pedoe, H. et al , 2007) (Tunstall-Pedoe, H. and Woodward, M., 2005). Concern has been expressed that a major programme designed to increase treatment of those at highest risk of CVD may increase social inequalities in health by undertreatment in the most deprived sections of society and overtreatment in the most affluent (Brindle, P., McConnachie, A., Upton, M. N. et al , 2005).

4.3.5 QRISK Narrative

During the last phase of the development of the guideline a new CVD risk score, QRISK, has been derived and validated using data from a UK primary care population (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007). Data were retrieved from the QRESEARCH database (www.qresearch.org), a large electronic database representative of primary care, and containing the health records of 10 million patients over a 17 year period from 529 general practices using the EMIS computer system. QRESEARCH contains area measures of ethnicity and also deprivation (Townsend score) based on the 2001 UK census, and linked to every patient's record. Information from two thirds of the QRESEARCH database was used for modelling dataset and the remaining third was used for validation dataset. An open cohort of patients aged 35 to 74 years at the date of study entry was identified that was drawn from patients registered from 1 January 1995 to 1 April 2007. The following patient groups were excluded; those with diabetes or CVD before their entry date into the database, temporary residents or those

with interrupted periods of registration at the practices and 4% of patients that did not have a valid postcode ethnicity score (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007).

The primary outcome was the first recorded diagnosis of CVD (including MI, CHD, stroke and transient ischaemic attack) on the general practitioners clinical computer system, either before or at death occurring between 1 January 1995 and 1 April 2007. The following risk factors were included in the analysis using the closest to the entry date to the cohort for each patient and imputing missing values when necessary; age (in single years), sex, smoking status (current smoker, non smoker-including former smoker), systolic blood pressure (continuous), ratio of total serum cholesterol to high density lipoprotein levels (continuous), left ventricular hypertrophy recorded on clinical records (yes or no), body mass index (continuous), family history of CVD in first degree relative aged less than 60 years (yes or no), body mass index (continuous), Townsend deprivation score, percentage of South Asian residents at output areas, current prescription of at least one antihypertensive (yes or no). A Cox proportional hazard model was used to estimate the coefficients associated with each potential risk factor for the first ever recorded diagnosis of CVD for men and women separately. The variables to be included in the model were specified a priori. Models were compared using the Bays information criterion (a likelihood measure which in lower values indicate better fit, and in which a penalty is paid for increasing variables). The strength of the association between one unit increases in each continuous risk factor was examined, and categories for other variables such as smoking compared with non-smoking were compared. The proportional hazards model's assumptions were tested for any non-linear relation between continuous independent variables and the outcome. Interactions between systolic blood pressure and antihypertensive treatment and also between smoking and deprivation were examined. The log of the hazard ratios for each of the risk factors (the coefficients from the Cox regression) from the model were used as weights for the new CVD risk equation. An estimate of each patient's probability of experiencing a CV event was made by combining these weights, the characteristics of the patient, and also using the baseline survivor

function for all participants. The baseline survivor function was estimated from the Cox regression model centred on the means of continuous risk factors, and the value for 10 year follow-up was extracted (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007).

The performance of the risk equation in the derivation dataset (QRISK score) was tested in the validation dataset by calculating the 10 year estimated CVD risk for each patient in the dataset. Missing values for continuous variables were replaced with mean values obtained from the derivation dataset by five-year age-sex bands, and assuming patients were non smokers if status was not recorded. Calibration (the degree of accuracy) was assessed by calculating the mean predicted risk of CVD at 10 years and the observed risk at 10 years obtained using the 10 year Kaplan-Meier estimate. The ratio of the predicted to the observed CVD risk for patients was then compared in patients in the validation cohort in each tenth of predicted risk. The predicted and observed risks were also compared for mean and women by age band and fifth of the Townsend score. Discrimination was assessed by receiver operated curve, and also by the R^2 and D^2 statistics (measures of discrimination and explained variation for survival models). The performance of QRISK was compared to the Framingham and ASSIGN equation (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007).

There were 478 UK practices that met the study inclusion criteria, 318 practices were randomly assigned to the derivation dataset (total patient number aged 35 to 74 years = 1 283 174, 50.4% women) and 160 practices to the validation dataset (total patient number aged 35 to 74 years = 614 553, 50.3% women). In the derivation dataset there were 65 671 incident cases of CVD and these were higher in men than women. The median follow up was 6.5 years. The 10 year observed risk of a CV event in women was 6.69% (95%CI 6.61% to 6.78%), and in men was 9.46% (95%CI 9.36% to 9.56%). In the validation dataset, the 10 year observed risk of a CV event in women was 6.60% (95%CI 6.48% to 6.72%), and in men was 9.46% (95%CI 9.14% to 9.43%). The final Cox regression model used in the study included the logarithm of age, ratio of serum cholesterol to HDL cholesterol, systolic blood

pressure, body mass index, family history of premature CHD, smoking status, Townsend deprivation score, and the use of at least one blood pressure treatment. The final model also included an interaction term between systolic pressure and blood pressure treatment. Left ventricular hypertrophy and the area measure of ethnicity were omitted. Hazard ratios for the final Cox regression analysis showed in the risk of CVD was increased with increasing age, body mass index and Townsend deprivation score. The risk was higher in patients who smoked, had a family history of CVD, and were receiving antihypertensive therapy. The hazard ratio for the ratio of total cholesterol to HDL cholesterol was just above and close to one, but it had been decided to include this factor a priori (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007).

From the calibration and discrimination modelling, the Framingham equation over-predicted risk at 10 years by 35%, ASSIGN by 36% and QRISK by 0.4%. All three equations tend to over predict risk in the lowest three tenths of the 10 years, the greatest over prediction occurred with ASSIGN, followed by Framingham and then QRISK. The receiver operator curve (ROC) statistic indicated that the final QRISK score had at least as good as, if not slightly better discrimination than the Framingham and ASSIGN equations. The R^2 statistics (standard error) for QRISK, Framingham and ASSIGN for women were; 36.4% (0.43), 31.7% (0.44) and 34.1% (0.43), respectively. The D^2 statistics (standard error) for QRISK, Framingham and ASIGN for men were; 33.3% (0.39), 29.1% (0.38) and 30.5% (0.38), respectively. Comparison of the proportion of patients with a CVD risk score ≥ 20% by Townsend fifths and sex for the three risk prediction scores found that the biggest difference was observed in women. QRISK predicted 9.8% of women aged 35 to 74 years from the most deprived fifth to be at high risk compared with 3.0% of women from the most affluent fifth. The corresponding values for the Framingham equation were 6.3% (most deprived) and 4.6% (most affluent). QRISK predicted 12.6% of men from the most deprived areas to be at high risk compared with 9.6% of those from the most affluent areas. The values for the Framingham equation were 19.5% (most deprived) and 20.5% (most affluent). Overall, QRISK predicted 8.5% of patients aged 35 to 74 years to be at high

risk compared with 12.8% for the Framingham equation and 14.0% for ASSIGN. Using QRISK, 34.5% of women and 72.9% of men would be at high risk compared with 24.1% and 86.0% using the Framingham equation (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007).

The performance of the QRISK score for predicting CVD risk was assessed in a second medical records database; The Health Improvement Network (THIN). This new electronic database contains records from general practices, some of which have or continue to participate in the General Practice Research Database (GPRD) and others that have never participated in the in GPRD. Hippisley-Cox *et al* identified the second cohort of patients from the THIN database, with the same inclusion and exclusion criteria as that for the original study (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2008), registered between 1 January 1995 and 31 March 2006. A Framingham score and QRISK score was generated for each individual patient in the THIN cohort and also the validation QRISK cohort. Hippisley-Cox *et al* used a revised equation for QRISK that had taken account of improvements in the method for multiple imputation of missing data in which additional variables (including the outcome variable) were included in the imputation model. The equation now excluded patients taking statins at baseline, and the results then differed from those previously reported (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2008).

There were 1 072 800 patients in the THIN cohort that were analysed (529 813 men (49.39%)). The corresponding cohort on QRESEARCH had 607 733 patients. The baseline characteristics were similar for THIN and QRESEARCH for age, sex, risk factors and medication, however, the family history of premature CHD was substantially lower in THIN than QRESEARCH (3.5% in males in THIN versus 9.2% in males in QRESEARCH). The Framingham equation over predicted risk by 28% in the THIN cohort while, QRISK under predicted by 10%. QRISK performed better than Framingham for the discrimination and calibration statistics (receiver operator curve statistic, R^2 statistic, D^2 statistic). The validation statistics for both QRISK and

Framingham were similar in the THIN cohort and the QRESEARCH cohort (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2008)

4.3.6 Cost effectiveness Narrative

There is no cost effectiveness evidence regarding the choice of tool. Refer to Section 4.2.3 of the full guideline.

4.3.7 Evidence to Recommendations

When the guideline started, the Framingham equation was the dominant method of calculating risk. Early in the development the GDG discussed the limitations of Framingham equation including:

- The tendency of Framingham equation to over estimate risk in modern European populations
- The tendency of Framingham equation to under-estimate risk in people from deprived backgrounds
- The difficulties in using Framingham in clinical practice when patients may already be on treatment
- Difficulties in adjusting Framingham for additional known risk factors such as a family history of CHD,
- Framingham equation being based on a fixed population with baseline data collected in the late 1960s and 1970s.

The GDG examined the existing literature on adjustments to Framingham and made recommendations on how the Framingham equation could be adjusted to the UK population.

In the later stages of development of the guideline the GDG became aware of the development of the QRISK equation and invited the principal investigator to attend a GDG meeting and present the preliminary findings. The GDG recognized the potential of a risk score developed in the UK population but only had preliminary data available to them. Two members of the GDG declared an interest in this area as researchers involved in the development of QRISK and were treated as experts for this discussion and any other

discussions on choice of risk score. They left the room and were not involved in decisions on choice of risk score.

Following consultation, the GDG considered stakeholder comments on the draft guideline, the first paper describing QRISK (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007) and the rapid responses to that paper including authors reply (http://www.bmj.com/cgi/content/short/bmj.39261.471806.55v1). The GDG also had access at this time to a second unpublished paper validating QRISK and addressing many of the criticisms in the original paper. The second paper is now published (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2008)

The performance of QRISK in this primary care population was better than the Framingham equation across each statistical measure. It reclassified a greater proportion of people from deprived backgrounds as being at high risk, relative to Framingham, as it took into account the increased risk associated with social deprivation. It appeared to address many of the limitations of Framingham because;

- in addition to standard risk factors QRISK includes variables relating to
  - Social deprivation
  - Being on BP treatment
  - Having a family history of CHD
  - Body Mass Index
- QRISK can be regularly updated and so keep up with secular changes in CVD incidence
- QRISK uses current primary care data to derive a risk score in the population in which it is to be used. i.e. UK primary care.

At the time of this meeting (September 2007) the GDG had two main concerns about recommending QRISK:

1 The GDG did not have the technical skills to assess the appropriateness and accuracy of the advanced statistical techniques (i.e. multiple imputation) employed.

2 Only one paper (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2007) had been published and subject to scientific review. This process had revealed some problems with the first equation. The subsequent paper detailing the corrections and adjustments {Hippisley-Cox, 2008 7945 /id} had not been published and subject to peer review and comment. Because of these concerns, the GDG (excluding the two researchers who left the room) felt unanimously that they were not able to recommend QRISK on the basis of the evidence available to them. They recommended to the Institute that either expert technical opinion be sought or that the guideline be published but might need early review.

As the Institute did not wish to update a guideline so soon after publication, it was agreed with the GDG that publication be delayed while independent expert opinion was sought. With the agreement of the GDG, the Institute sought advice from experts independent of the groups that had derived either QRISK or modified the Framingham equations or guidelines that advocate them. Advice was sought from a:

- Biostatistician:- Professor Doug Altman
- Epidemiologist: - Professor Sir Richard Peto FRS
- Expert in Cardiovascular Risk Estimation: Professor Rod Jackson

Their reviews are attached as an appendix.

The GDG reconvened in January 2008 to discuss the now published QRISK paper (Hippisley-Cox, J., Coupland, C., Vinogradova, Y. et al , 2008) and the independent reviews. The GDG discussed the independent reviews and sought clarification of some points from the two researchers. The GDG addressed methods for dealing with missing data, calibration and discrimination statistics for QRISK and the applicability and use of QRISK in different clinical settings. The GDG (excluding the two researchers who left

the room) unanimously agreed that QRISK should be recommended. The GDG agreed that the recommendation of QRISK will also allow the score to be improved with the potential to include other variables and outcomes of interest.

The GDG had some outstanding concerns:

1) The calculation of the additional risk of some ethnic groups, in particular those of south Asian background.

The QRISK equation does not include a variable for ethnicity, but does include a variable for deprivation and family history. The previous recommended increase of a factor of 1.4 in risk for South Asian males when using the Framingham equation would overestimate the risk using the QRISK equation. As there is no information currently available on what, if any, increase would be appropriate for ethnicity, the GDG decided not to include any adjustment This has been recommended for further research.

2) The management of patients who had previously been assessed with the Framingham equation and were currently on treatment. The GDG regarded it as inappropriate for a patient currently on treatment to be reassessed with the possibility of the treatment being stopped. The GDG agreed that patients already on treatment should not be reassessed using QRISK.

3) Accessibility of QRISK

The view of the GDG is that QRISK must be freely available for incorporation into primary care management software and to secondary care clinicians for use in hospital. The GDG will ask for a guarantee from the developers that the algorithms will be freely available from their website prior to publication.

4) Updating the algorithms

A major advantage of QRISK is that it can be updated to, for example, reflect changes in the UK population, or to include more variables such as ethnicity. However there must be strict version control, therefore the GDG recommends

that NICE work with developers to co-ordinate updates in QRISK with the publication of updates of the guideline.

Reference List

(1) Joint British recommendations on prevention of coronary heart disease in clinical practice. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, endorsed by the British Diabetic Association. Heart. 1998; 80 Suppl 2 :S1-29.

(2) Joint British recommendations on prevention of coronary heart disease in clinical practice: summary. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, British Diabetic Association. BMJ. 2000; 320 (7236) :705-708.

(3) 1996 National Heart Foundation clinical guidelines for the assessment and management of dyslipidaemia. Dyslipidaemia Advisory Group on behalf of the scientific committee of the National Heart Foundation of New Zealand. N Z Med J. 1996; 109 (1024) :224-231.

(4) Aarabi M, Jackson PR. Predicting coronary risk in UK South Asians: an adjustment method for Framingham-based tools. Eur J Cardiovasc Prev Rehabil. 2005; 12 (1) :46-51.

(5) Anderson KM. Cardiovascular disease risk profiles. Am Heart J. 1991; 121 (1 part 2) :293-298.

(6) Bastuji-Garin S, Deverly A, Moyse D, Castaigne A et al. The Framingham prediction rule is not valid in a European population of treated hypertensive patients. J Hypertens. 2002; 20 (10) :1937-1980.

(7) Beswick AD, Brindle P, Fahey T, Ebrahim S. A systematic review of risk scoring methods and clinical decision aids used in the primary prevention of coronary heart disease.

(8) Bhopal R, Fischbacher C, Vartiainen E, Unwin N et al. Predicted and observed cardiovascular disease in South Asians: Application of FINRISK, Framingham and SCORE models to Newcastle Heart Project data. J Public Health. 2005; 27 (1) :93-100.

(9) Brindle P, Emberson J, Lampe F, Walker M et al. Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. BMJ. 2003; 327 (7426) :1267-1272.

(10) Brindle P, May M, Gill P, Cappuccio F et al. Primary prevention of cardiovascular disease: a web-based risk score for seven British black and minority ethnic groups. Heart. 2006; 92 (11) :1595-1602.

(11) Brindle P, McConnachie A, Upton MN, Hart CL et al. The accuracy of the Framingham risk-score in different socio-economic groups: a prospective study. 2005.

(12) Brindle PM, Beswick AD, Fahey T, Ebrahim SB. The accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: A systematic review. Heart. 2006; 92 (12) :1752-1759.

(13) Cappuccio FP, Oakeshott P, Strazzullo P, Kerry SM et al. Application of Framingham risk estimates to ethnic minorities in United Kingdom and implications for primary prevention of heart disease in general practice: cross sectional population based study. BMJ. 2002; 325 (7375) :1271-1274.

(14) Carlsson R, Lindberg G, Westin L, Israelsson B. Serum lipids four weeks after acute myocardial infarction are a valid basis for lipid lowering intervention in patients receiving thrombolysis. Br Heart J. 1995; 74 (1) :18-20.

(15) Conroy RM, Pyorala K, Fitzgerald AP, Sans S et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J. 2003; 24 (11) :987-1003.

(16) Department of Health. Tackling Health Inequalities: a programme for action. London: Department of Health. 2003

(17) Empana JP, Ducimetiere P, Arveiler D, Ferrieres J et al. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME Study. Eur Heart J. 2003; 24 (21) :1903-1911.

(18) Gill PS, Kai J, Bhopal RS, Wild S. Health care needs assessment: black and minority ethnic groups. In: Raferty J, Stevens A, Mant J, editors. Health Care Needs Assessment: the epidemiologically based needs assessment reviews. Abingdon: Radcliffe Medical Press; 2007.

(19) Greenland P, La Bree L, Azen SP, Doherty TM et al. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. JAMA. 2004; 291 (2) :210-215.

(20) Grover SA, Coupal L, Hu XP. Identifying adults at increased risk of coronary disease. How well do the current cholesterol guidelines work? JAMA. 1995; 274 (10) :801-806.

(21) Grover SA, Dorais M, Coupal L. Improving the prediction of cardiovascular risk: interaction between LDL and HDL cholesterol. Epidemiology. 2003; 14 (3) :315-320.

(22) Hall LM, Jung RT, Leese GP. Controlled trial of effect of documented cardiovascular risk scores on prescribing. BMJ. 2003; 326 (7383) :251-252.

(23) Hanon O, Franconi G, Mourad JJ, Baleydier A et al. L'estimation du risque cardiovasculaire chez des patients hypertendus ne modifie pas la prise en charg de l'HTA. Arch Mal Coeur Vaiss. 2000; 93 (8) :943-947.

(24) Haq IU, Jackson PR, Yeo WW, Ramsay LE. Sheffield risk and treatment table for cholesterol lowering for primary prevention of coronary heart disease. Lancet. 1995; 346 (8988) :1467-1471.

(25) Hense HW, Schulte H, Lowel H, Assmann G et al. Framingham risk function overestimates risk of coronary heart disease in men and women from Germany--results from the MONICA Augsburg and the PROCAM cohorts. Eur Heart J. 2003; 24 (10) :937-945.

(26) Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J et al. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. Heart. 2008; 94 (1) :34-39.

(27) Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J et al. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. BMJ. 2007; 335 (7611) :136.

(28) Jackson R. Updated New Zealand cardiovascular disease risk-benefit prediction guide. BMJ. 2000; 320 (7236) :709-710.

(29) Lloyd-Jones DM, Nam BH, D'Agostino RB, Sr., Levy D et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: a prospective study of parents and offspring. JAMA. 2004; 291 (18) :2204-2211.

(30) Marshall T. The use of cardiovascular risk factor information in practice databases: Making the best of patient data. Br J Gen Pract. 2006; 56 (529) :600-605.

(31) Marshall T, Rouse A. Resource implications and health benefits of primary prevention strategies for cardiovascular disease in people aged 30 to 74: mathematical modelling study. BMJ. 2002; 325 (7357) :197.

(32) McCormack JP, Levine M, Rangno RE. Primary prevention of heart disease and stroke: a simplified approach to estimating risk of events and making drug treatment decisions. CMAJ: Canadian Medical Association Journal. 1997; 157 (4) :422-428.

(33) McElduff P, Lyratzopoulos G, Edwards R, Heller RF et al. Will changes in primary care improve health outcomes? Modelling the impact of financial incentives introduced to improve quality of care in the UK. Qual Saf Health Care. 2004; 13 (3) :191-197.

(34) McLeod AJ, Armitage M. Use of statins. But New Zealand tables are better. BMJ. 1998; 317 (7156) :474.

(35) Nam BH, Kannel WB, D'Agostino RB. Search for an optimal atherogenic lipid risk profile: from the Framingham Study. Am J Cardiol. 2006; 97 (3) :372-375.

(36) Nazir DJ, Roberts RS, Hill SA, McQueen MJ. Monthly intra-individual variation in lipids over a 1-year period in 22 normal subjects. Clin Biochem. 1999; 32 (5) :381-389.

(37) Nilsson PM, Nilsson JA, Berglund G. Family burden of cardiovascular mortality: risk implications for offspring in a national register linkage study based upon the Malmo Preventive Project. J Intern Med. 2004; 255 (2) :229-235.

(38) Primatesta P, Poulter NR. Levels of dyslipidaemia and improvement in its management in England: results from the Health Survey for England 2003. Clin Endocrinol (Oxf). 2006; 64 (3) :292-298.

(39) Primatesta P, Poulter NR. Lipid levels and the use of lipid-lowering agents in England and Scotland. Eur J Cardiovasc Prev Rehabil. 2004; 11 (6) :484-488.

(40) Quirke TP, Gill PS, Mant JW, Allan TF. The applicability of the Framingham coronary heart disease prediction function to black and minority ethnic groups in the UK. Heart. 2003; 89 (7) :785-786.

(41) Ramsay LE, Haq IU, Jackson PR, Yeo WW et al. Targeting lipid-lowering drug therapy for primary prevention of coronary disease: an updated Sheffield table. Lancet. 1996; 348 (9024) :387-388.

(42) Ryder RE, Hayes TM, Mulligan IP, Kingswood JC et al. How soon after myocardial infarction should plasma lipid values be assessed? Br Med J. 1984; 289 (6459) :1651-1653.

(43) Sesso HD, Lee IM, Gaziano JM, Rexrode KM et al. Maternal and paternal history of myocardial infarction and risk of cardiovascular disease in men and women. Circulation. 2001; 104 (4) :393-398.

(44) Thompson SG, Pocock SJ. The variability of serum cholesterol measurements: implications for screening and monitoring. J Clin Epidemiol. 1990; 43 (8) :783-789.

(45) Tunstall-Pedoe H, Woodward M. By neglecting deprivation cardiovascular risk scoring will exacerbate social gradients in disease. Heart. 2005; 92 (3) :307-310.

(46) Wallis EJ, Ramsay LE, Ul H, Ghahramani P et al. Coronary and cardiovascular risk estimation for primary prevention: validation of a new Sheffield table in the 1995 Scottish health survey population. BMJ. 2000; 320 (7236) :671-676.

(47) Warnick GR. Measurement of cholesterol and other lipoprotein constituents in the clinical laboratory. Clin Chem Lab Med. 2000; 38 (4) :287-300.

(48) Westgard JO, Darcy T. The truth about quality: medical usefulness and analytical reliability of laboratory tests. Clin Chim Acta. 2004; 346 (1) :3-11.

(49) White C, von Galen F, Chow YH. Trends in social class differences in mortality by cause. Health Stat Q. 2003;(20) :25-30.

(50) Wilson PWF, D'Agostino RB, Levy D, Belanger A et al. Special Report: Prediction of Coronary Heart Disease Using Risk Factor Categories. Circulation. 1998; 97 :1837-1847.

(51) Wood D, De B, Faergeman O, Graham I et al. Prevention of coronary heart disease in clinical practice: recommendations of the Second Joint Task Force of European and other Societies on Coronary Prevention. Atherosclerosis. 1998; 140 (2) :199-270.

(52) Wood D, Wray R, Poulter N, Williams B et al. JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. Heart. 2005; 91 (supp 5) :v1-52.

(53) Woodward M, Brindle P, Tunstall-Pedoe H, SIGN group. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart. 2007; 93 (2) :172-176.

Appendix: Expert reviews

The following reviews give views on QRISK of three experts commissioned by NICE. These are included for information only and are not for comment.

Review from Professor Doug Altman

Notes on QRISK development and validation studies

Doug Altman, 17 December 2007

Summary

- I believe that the development of the QRISK score, after revision, was based on appropriate statistical methods and that the validation studies were also performed appropriately.
- The QRISK score performed well in two validation samples.
- Some concerns about the quality of some of the data would not affect the observed performance. Also the similarity of the derivation and validation cohorts has been noted, but these do represent the population on whom the score would be used.

Specific comments

1. The authors developed a prognostic model to provide a new cardiovascular disease risk score, using data from a database derived from GP consultations.
2. The sample size was massive. However, large sample size cannot compensate for any weaknesses in the data. There seem to be legitimate concerns about the quality of some of the data such as smoking status [2]. Imprecise data would reduce performance compared to good data, but this possibility does not weaken the *observed* findings.
3. The statistical methods of model derivation were sound, including careful analysis of continuous predictors. Several models were produced – I have focused only on the authors' preferred model A.
4. As there was a lot of missing data for several variables the authors used multiple imputation (MI). There is an increasingly widely held view that imputation of missing values yields less biased results compared to

complete case analysis. The use of this approach here is supported by the observed associations between missingness and some of the prognostic variables. In particular, total/HDL cholesterol ratio was absent for two thirds of cases, and missingness was associated with a poorer outcome.

5. The surprising failure to detect total/HDL cholesterol ratio as important in the model was noted by several commentators [2].
6. The authors' helpful responses to various comments and criticisms [3] included corrected models but did not fully specify what had been changed – that information has now been made clear in a technical supplement [5]. Unfortunately, as that document makes clear, the authors' original implementation of MI was faulty. In particular, in their updated analysis they rectified the important error of omitting the outcome (dead or not) from the imputation model. (In the revised analyses they also additionally omitted a few patients on statins at baseline. I agree with this change, which would have minimal impact.)
7. There was very little impact of the changes to the imputation procedure on the performance of QRISK in either the derivation or validation data sets. However, the proportion of patients at high risk (>20%) was slightly reduced in the revised analysis.
8. The original paper [1] did not include some key information that has now been provided in the technical supplement [5]. For example, the BMJ paper did not specify how many of the patients had complete data; this is now revealed: "24% of women and 22% of men had complete data for all risk factors used in the Cox regression model" [5].
9. The main analysis used Cox regression, which does not give simple predictions of proportions surviving. To compare observed and expected survival at 10 years the authors used the baseline hazard with all continuous variables set at the mean (but they don't say anything about non-continuous variables).
10. The median follow up was 6.5 years, range 1-12. Thus taking 10 years for predictions is perhaps unwise. The paper doesn't say how many patients were in fact followed for 10+ years – I suspect it would be a small proportion of the whole sample. Because of the huge sample size, this issue may not be of great consequence.
11. Separate models were fitted for men and women – these are very similar and I am not clear if there is much gain in having separate models.

12. Two validation exercises were performed. As part of the original study the authors reserved one third of practices for validation [1]. In addition, they assessed performance on data from THIN, a rather similar database also including data from UK general practices [4].
13. Missing data in the validation data sets were imputed rather simply using data from the derivation cohort. The performance of the revised QRISK on the first validation data set was minimally changed after the corrected imputation. (In the reanalysis they also presented results after multiple imputation within the validation data set, which showed somewhat better performance – Appendix 3 in [5]).
14. The performance in the two validation data sets was extremely similar.
15. The performance of QRISK is good for such a predictor, with ROC area of about 0.78.

[1] Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* 2007:bmj.39261.471806.55.

[2] *BMJ* rapid responses (latest dated 31 Oct 2007)

[3] Hippisley-Cox J CC, Vinogradova Y, Robson J, May M, Brindle P. QRISK: Authors Response. http://www.bmj.com/cgi/eletters/335/7611/136: *British Medical Journal*, 2007.

[4] Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. The performance of the QRISK cardiovascular risk prediction algorithm in an external UK sample of patients from general practice: a validation study. *Heart* 2007:hrt.2007.134890.

[5] Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. QRISK Cardiovascular Disease Risk Prediction Algorithm – comparison of the revised and the original analyses. Technical Supplement 1. 01 November 2007.

Review from Professor Richard Peto

16 January 2008

My general comment on QRISK reflects the concerns of the 6 July 2007 BMJ editorial on it by Bonneux ("Cardiovascular risk models — the moral implications"). **Should** a 10-year vascular risk of over 20% really be the only measure of when to use prophylactic treatment? If so, virtually all apparently healthy men aged over 70 (and absolutely all those over 80) "should" be on treatment. The views of NICE on QRISK should be determined by exactly what use is to be made of it: my own view is that heart attacks at 60 should matter a lot more than heart attacks at 80, and I'd want the use in general practice of any mechanical risk calculation formulae to reflect this. Some hurried notes follow.

Best wishes,

Yours sincerely

Richard Peto

Richard Peto

Brief notes on QRISK follow on next page

Brief notes on QRISK

Framingham risk scores or QRISK risk scores? If cardiovascular risk is to be predicted for apparently healthy patients (with no history of vascular disease or diabetes) in general practices in England & Wales then it's better to do this on the basis of recent local data (eg, QRISK) rather than previous non-UK data (eg, Framingham).

Age, sex & other factors: The main risk predictors are simply **age and sex**, as is illustrated by table 11 of the 1 November 2007 technical supplement (section 10.6, page 10-21). (Hence, one convenient and understandable way to summarise information on other factors such as total/HDL cholesterol, systolic blood pressure [SBP], smoking or social class might be to state approximately how many years older or younger a particular measurement makes you seem, in comparison with having no information about that factor.) Hence, any tables of calibration and discrimination for QRISK (eg, pp.10-18 et seq) should mainly address the question of how much QRISK **adds** to an optimal score based only on sex and age.

Treatment of missing values: It disturbs me that, on pages 10-12 to 10-14, the 10-year risks are twice as big for women with missing smoking, BMI, SBP or cholesterol than for women with these things measured.

Regression coefficients: It also disturbs me that smoking seems to carry a relative risk of only 1.5 (table 7 on p.10-17) for cardiovascular disease, when, for example, the prospective Million Women Study finds a smoker versus never-smoker relative risk of about 4 for vascular mortality. Are many non-smokers wrongly classified?

Likewise, prospective studies have, collectively, shown that a 20 mmHg difference in usual SBP is associated with 1/3 less vascular mortality; why is the effect so small in Table 7? (I suppose it's because [i] only the measured SBP is used, [ii] treatment of BP enters as a positive risk factor and [iii] an SBP/treatment interaction is fitted, but it's still potentially misleading.)

The BMJ QRISK article in July got the regression coefficient completely wrong for total/HDL cholesterol, but the current QRISK revision corrects the error.

Imputed values: I still don't really trust this procedure, but that may simply be because I haven't gone deeply enough into it to know exactly what was done.

Review from Professor Rod Jackson

A critique of QRISK versus Framingham CVD risk prediction scores for NICE.

Professor Rod Jackson (2 December 2007).

Introduction:

I have based this critique primarily on the two published QRISK papers (BMJ 2007 and Heart 2007 on-line), the authors' letter to the BMJ (2007) responding to comments on their BMJ paper, and on multiple email communications with the QRISK authors over the last two months (October – November 2007). Fortuitously, several months ago I was asked by the editor of Heart to write an editorial on the QRISK validation paper. The proofs of this editorial are attached and it will be published in paper form, along with the QRISK paper, in Heart in January 2008 (an on-line version of the QRISK validation has already been published). (Editors' note: this editorial has now been published Jackson R (2008) Cardiovascular risk prediction: are we there yet? *Heart* 94: 1–3)

This editorial contains the bulk of my critique and I recommend that it is read as an introduction to this critique. It also describes the context of risk assessment and identifies the key questions for guideline developers and clinicians. Below I summarise the main issues in my critique and add some new information I have received about the QRISK score from the authors, since my editorial was submitted. The comments below are made on the assumption that the reader has already read the attached editorial.

Critique:

The key characteristics of a high quality clinical risk prediction tool are that it: i. is well calibrated; ii. is able to reasonably discriminate between those people

who will develop the condition predicted from those who will not develop the condition, in a defined period; and iii. is able to be effectively and cost-effectively implemented in the clinical context for which it has been developed.

i. Calibration.

This is the ability of a tool to predict, at a group level, a similar level of risk (e.g. 15-20% 10 year risk) as the observed risk in that group of patients, in the appropriate clinical population. In both the BMJ and Heart validation studies, QRISK clearly excels on calibration compared with Framingham. For example, in the independent validation study described in the Heart publication, as I discuss in my editorial, the predicted QRISK scores was about 10% lower than the observed risk for both men and women while for Framingham it was 16% higher for women and 28% higher for men. While calibration is a very important component of risk prediction, it is also the easiest to adjust and the Framingham group have described a recalibration process that has been successfully used to recalibrate the Framingham equation for a range of populations (reference 15 in my editorial).

ii. Discrimination.

This is the ability of a risk prediction tool to differentiate between patients who will develop the predicted condition and those who won't develop it, in a specified time period. There are two main measures of discrimination that I have called summary (or global) discrimination and specific (or threshold) discrimination. The former measure (i.e. global) assesses discrimination over the whole range of possible prediction thresholds and is best shown visually by a receiver operator characteristic (ROC) curve. There is also a range of associated statistical scores for global discrimination and QRISK consistently scores better than Framingham, although it is difficult to determine the clinical significance of these differences, because in clinical practice usually only one threshold of predicted risk is used to inform treatment decisions. The latter

measure of discrimination (i.e. threshold) is more clinically relevant and at the time of writing my editorial, I did not have information on threshold discrimination for QRISK compared to Framingham. However I have since received this information from the QRISK authors for the current recommended treatment threshold for statins in the UK (i.e. 20% 10 year CVD risk) and this is quoted below.

'Of the 7.99% of patients with a QRISK score of 20% or more on the QResearch validation cohort, there were 7,555 of patients with CVD events over ten years. This represents 26.8% of the total number of patients with events (n=28,168). Looking at the 7.99% of patients with the highest Framingham scores, there were 7,019 patients with CVD events over 10 years (ie 24.9% of the number of patients with events).'

So, the QRISK tool is only slightly more discriminating than Framingham at the recommended treatment threshold – the 8% of patients at the highest predicted risk account for almost 27% of all events using QRISK and almost 25% using Framingham. Of note, this improvement, albeit small, is probably accounted for by the additional variables included in the QRISK equation but the reason there is only a small improvement is probably related to the large amount of missing data in the QRISK cohort. Another explanation for the quite small improvement in discrimination is that follow-up was censured if patients left the practice, so it is possible that the true threshold discrimination of QRISK is better than documented above.

iii. Application.

For a risk prediction tool to be useful, it must be applicable in routine practice. One important difference between QRISK and Framingham is that the QRISK score includes additional variables (i.e. BMI, family history of CVD, a deprivation index and current antihypertensive drug treatment). As discussed above these additional variables probably account for the slightly better discrimination of QRISK compared to Framingham, however adding these

variables comes at a cost. First it will require more time for the clinician to generate these variables but more importantly, the QRISK score will most likely require a computer for the calculation whereas the Framingham score, with fewer variables, can be estimated using a paper chart. While on the one hand this may be considered a weakness of QRISK, on the other it will ultimately become a strength because clinicians will be motivated to document all the required variables in order to generate a risk estimate. Currently the main weakness of QRISK is the missing risk factor data in the cohort (including 60-70% of lipid levels). I do not know if the deprivation score is readily available in GP records but I assume it is relatively easy to derive. If so, then all the additional variables required for the QRISK score are reasonable to expect to be documented in a GPs record.

Other issues.

On initially reading the QRISK papers I had some concerns about missing data. I have touched on the issue of missing risk data above which will become less of a problem if the QRISK score is implemented electronically. I am also reassured by the improved lipid level imputations described in the letter to the BMJ, which has led to QRISK having regression coefficients for lipids that are more in line with the international literature. I have also asked the QRISK authors about validation of non-fatal outcome data and they have provided indirect evidence that I found reassuring.

One other issue that is yet to be resolved is how to predict risk in patients on statins. These patients were excluded from QRISK, and while they were a relatively small proportion of the QRISK cohort, they are a growing group in practice. One possible response to this problem is to assume those already on statins have been identified as at high risk, but I think it is likely that for people without a history of CVD, the lipid level and not the risk level, remains the main determinant of treatment. To address this issue, there will need to be further developments of the QRISK score over time, however this should be much easier if QRISK is used as part of a computerised system.

As an aside, in New Zealand we predict risk in diabetics using the same Framingham equation that we use for all patients. I note that in the UK diabetics are treated as a separate group. I believe this is a mistake as diabetics are actual part of the same group of 'high CVD risk' patients and the most important and effective interventions are those addressing standard CVD risk factors rather than glycaemia. I understand it would have been possible for QRISK to include diabetes in the same score and would suggest this is considered in future.

Conclusions:

In my opinion there are sufficient improvements in the QRICK score compared with the Framingham-based risk scores to justify the use of QRISK as the most appropriate CVD risk assessment tool for the UK. The key implication of this recommendation however is that it will require risk prediction to be done electronically. As discussed, I believe this would be a positive step, because it will lead to significant improvements in the completeness of CVD risk factor documentation. Not only is this of clinical relevance in terms of the quality of clinical records but it will also enhance the development of more accurate risk prediction tools in future.