Reviewer's report

Title: Prediction of cardiovascular risk using default risk factor values compared to clinically estimated cardiovascular risk

Version: 1 Date: 25 October 2007

Reviewer: Islay Gemmell

Reviewer's report:

General
This paper describes a method of identifying patients who may have a cardiovascular risk of 20% over 10 years using routine data available from patient records. The author suggests that this initial assessment of risk can be used to identify patients who should be prioritised to receive a comprehensive risk factor assessment that requires patient time, staff time and laboratory tests. The method uses prior (default) estimates of risk factors such as blood pressure and cholesterol levels where these are not recorded on the patients’ records. The default values are calculated using population level estimates of risk factors derived from the Health Survey for England 1998.

The author calculated a ‘true’ CVD risk for patients and several other estimates of CVD risk based on different levels of availability of data that might exist in a typical practice. ROC curves were used to compare these various methods of estimating risk with the ‘true’ risk.

One the whole I think this is an important paper and I think it demonstrates a method of selecting patients for CVD risk assessment that would be beneficial to patients and practices alike. I think this methodology is a valuable and practical addition to the developments in CVD risk assessment.

-----------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

None

-----------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

In the text the author refers to six estimates of risk, these appear to have been calculated as below:

True – from HSE 2003 using Framingham
Clinically estimated – from HSE 2003 using Framingham + some additional variability to reflect the clinical setting (is this correct?)
Semi-complete data - age, gender, antihypertensive status, diabetic status, smoking status and 1 measure of BP
Minimum data – age, gender, antihypertensive status, diabetic status
Diabetic status and hypertensive status
Age alone

There were several instances thereafter when the description of the methods of estimating risk were inconsistent with those described in the methods section (above).

Firstly, the labels on the figure do not correspond directly with the descriptions given in the methods section. Antihypertensive status is not mentioned in the definition of the minimum data estimate and smoking status is not included in the definition of the semi-complete data estimate.

Secondly, in the 2nd to last sentence on page 8 the author does not include the age only curve in his description of what is contained in the figure.

Thirdly, in the second paragraph of the results section the author refers to the curve based on age, sex and diabetic status alone but I thought that this curve also included antihypertensive status as described on pages 7 and 8.

I think it would improve the paper considerably if a simple label was given to each method of estimating risk and this label (rather than a description of the method) was used thereafter in the paper.

In the results section I was unclear as to how the percentages of the population who could be assessed were derived. I think this could be explained more clearly.

I found the description of how the default values of the risk factors was calculated somewhat difficult to understand. (paragraph 4 in methods). However on referring to the website I quickly understood how these were calculated. Is it feasible for the calculation of the default values to be described in the paper in the same way that they are described in the website or does this breach conditions of publication?

2nd to last sentence in the discussion needs a comma between diabetic status and blood pressure.

Discretionary Revisions (which the author can choose to ignore)

In this example I think that sensitivity (identifying true +ves) is more important than specificity (identifying true –ves) since the impact of false positives i.e. asking a patient to come for an assessment when they are not at high risk is of little consequence to the patient and as far as the practice is concerned the patient may have been selected for assessment under the old system. However we want this method to clearly identify those at high risk so they are not excluded
(or delayed) in having their risk assessment. Perhaps the author would like to comment on this in the discussion.

**What next?:** Accept after minor essential revisions

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests