Reviewer's report

Title: Description and validation of a Markov model of survival that uses Framingham risk factors and smoking status.

Version: 1 Date: 10 February 2004

Reviewer: Alan Montgomery

Reviewer's report:

General

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

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

Overall, I felt that the Methods and Results sections needed more detail in order for someone else with reasonable competency in decision modelling to be able to reproduce the analyses, and to understand the results presented. Specific points that require further clarification:

(1) The description of the model in the text and Figure 1 talk about ‘calculating’ or ‘estimating’ mortality due to different causes. In a Markov model these probabilities are defined by the modeller at the outset, using data from observational studies, routine statistics, randomised trials etc. The probabilities are entered into the model, and individuals progress through the model each time period (in this case one year). Probabilities are not ‘re-calculated’ during the model, though may be re-defined if they change as someone gets older. This may just require some re-phrasing, or there may indeed be some flaw in the model, but either way this should be clarified.

(2) As a further point, probabilities should strictly be sampled from a distribution during each time cycle, so that the model is truly stochastic rather than deterministic. A point estimate and characteristics of the distribution are defined (such as normal distribution with 95% CI), the probability of an event is sampled from this distribution on each cycle, and the uncertainty inherent in all estimates of event probabilities is built into the model. If the authors do not want to increase the sophistication of the model to this level, they should at the very least conduct some sensitivity analyses that demonstrate the impact (or perhaps lack of impact) of changing the probabilities used in the model. A table presenting the probabilities used in the base case scenario and any sensitivity analyses would be helpful. These should be annual probabilities, rather than relative, or any other representation of risk.

(3) It wasn’t clear from either the description of the model or Figure 1 whether someone is allowed to have more than one MI or stroke, and if so, which one is assumed to be fatal.

(4) Validation. Almost 1400 patients from the cohort are excluded from the analysis. Some of these will be >65 years, but excluding many other cases because of missing data may introduce bias into the analysis. This should be discussed in the ‘limitations’ section.

(5) As I understand it, Table 2 presents the proportions of individuals still alive after 20 years as estimated by the model and observed in the cohort, the difference between these proportions, a 95% confidence interval, and a p-value for the test of the null hypothesis that the difference equals zero. The final column shows the power of the given sample size to detect a 20% difference between the two proportions.
(a) By conducting repeated hypothesis tests, the authors increase the likelihood of finding a 'significant' difference by chance alone.

(b) I can’t find any reasoning behind conducting post-hoc power calculations based on a 20% difference between the estimated (model) and observed (cohort) 20 year survival. For agreement, if one is going to do such calculations at all, it makes sense to specify a difference beyond which one would not accept that the model makes a good estimate of survival, and I would have thought that this difference would be much smaller than 20%. For a smaller difference, say 5%, this is obviously going to drastically reduce the power of the study. However once the analyses have been done, there is really no need to do such post-hoc power calculations. Simply present the proportions in the model and cohort, difference and 95% CI, and possibly the p-values, and allow the reader to interpret these. The column containing relative differences isn’t very helpful and could be removed.

(6) Apart from a couple of categories with small sample sizes, the agreement between the model and observed survival strikes me as remarkably close. This re-iterates the need to make the description of the model and the sources of all the probabilities much clearer.

(7) Results paragraph 2 should be in the methods section. This also wasn’t completely clear to me. “… cases were sorted into ascending order of estimated survival at 20 years.” I thought that individuals were classified as either ‘alive’ or ‘dead’ after 20 years, so on what basis were they sorted and arranged into 40 groups? Figure 2 should make clear what each data point represents (ie what characteristics). Given that maximum survival can only be 1.0, the figure axis should be changed.

(8) Where did the p-values quoted in the Discussion section paragraph 1 come from? Also see point above about whether a 20% discrepancy between the model and cohort really represents acceptable agreement.

(9) Discussion page 11 and Figure 3. Is this a comparison of two 35 year old men, one with existing CVD and one without and stays CVD-free throughout his life? Or is it an estimate of treating a 35 year old man’s high blood pressure, high cholesterol and getting him to stop smoking? This basis of comparison should be clearer, should be described in the methods section, and results presented in the results section.

(10) Although this model makes assumptions about smoking behaviour, it is certainly possible to model quit/relapse rates and associated changes in risk in a more sophisticated model.

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

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

(11) The title of the paper specifically mentions smoking status. However the analysis is not particularly focussed on smoking behaviour.

(12) Although one could probably guess from the Background section, the aims of the study are not clearly stated.

(13) The key to the formula to calculate absolute risk in smokers (p6) contains definitions for NS and RN but these terms are not actually included in the formula.

(14) Is the proportion 0.803 (p7 para 4) a typing error? A 20% reduction from 0.754 would be 0.603. However consider removing this section (see above).

----------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

**Statistical review:** No

**Declaration of competing interests:**

None