Reviewer's report

Title: Predictors of mortality of patients newly diagnosed with clinical type 2 diabetes: a 5-year follow up study

Version: 2 Date: 1 March 2010

Reviewer: John Yudkin

Reviewer's report:

This reviewer was asked to comment on the revised manuscript in order to see whether the comments of Reviewers 1 and 3 had been adequately answered in the latest draft.

Olivarius et al have analyzed the predictors of mortality from a population-based sample of 1,323 persons newly diagnosed with clinical diabetes and age above 40 years. The main topic of the present manuscript was the evaluation of a simple questionnaire about the self-rated general health (SRH). SRH was evaluated with a single question: in general, how would you rate your health at present? The response categories were excellent, good, fair, poor and very poor. The authors concluded that patients who rated their health as less than excellent had increased 5-year mortality, even when biochemical, clinical and life-style variables were controlled for.

The authors have clearly put in an enormous amount of work to perform the analyses for this and previous drafts of the manuscript, which they clearly think is warranted by their conclusions. This reviewer feels that in general the authors have dealt with the anxieties of Reviewers 1 and 3, and these points will be considered below. There are, however, a number of separate problems, many of these being related to the methods and interpretations of the statistical analyses. As this reviewer has no statistical expertise, it is important that the comments of the previous reviewer 2 are re-requested, but there is an additional problem. This reviewer is probably around the median when considering the knowledge of the readers of BMC Endocrinology on statistical and epidemiological methods. This in turn means that things which confuse this reviewer may well confuse the audience, if the paper is accepted. Indeed I note that one of my main puzzles, the hazard ratios for age, were also not clear to the reviewer (2) with major statistical expertise. Thus much clearer explanation is called for. Additionally, while the majority of readers may be familiar with Information Criteria AIC and BIC and their significance, this is not something with which this reviewer is familiar. It might be useful to expand the explanation.

I will now itemize my concerns, starting with the major ones which have not been raised by other reviewers.

1. The main problem with the measure of self-rated health is likely to be confounding. The authors do not show the relationships of self-rated health with
any of the variables collected, and deal with the confounding issue by multivariate Cox analyses. The fact that inclusion of pre-existing disease substantially reduces the HRs, and the significance of the Wald test is reduced from P=0.007 to P=0.036, leaves open the possibility of residual confounding, bearing in mind that disease was self-reported. Furthermore the relationship of self-rated health with age is important, particularly bearing in mind that it is not clear to me whether chronological age is included in the analyses (see below).

2. The authors note that the likely reason that the expected relationships of risk factors with mortality is not seen may be because of the metabolic disequilibrium within a week or two of diagnosis of diabetes. But this metabolic disequilibrium may well also impact on self-rated health. The authors have explored self-rated health in people who have just been diagnosed with diabetes, and who are in the majority of cases likely to have had symptoms related to hyperglycaemia and polyuria for a variable period of time. It seems likely that self-rated health would therefore be affected by the degree of elevation of plasma glucose level, the renal threshold and by the duration of symptoms. These last two are unavailable, but diagnostic plasma glucose is.

3. The consideration of age in Table 1 and 3 is analysed as follows: ‘death intensity was represented as a function of patient age, multiplicatively affected by the characteristics.’ It is not clear to me how this was done. If this was adjusted for within the sample, then the effect of age should have disappeared. But it now seems that younger patients had a ‘relatively higher mortality than older patients (reviewer 2 point 6). This is likely to be incomprehensible to the average reader.

4. There are a number of variables in tables 1 and 3 where the HR from the Cox model differs greatly from an RR calculated from the total numbers. Thus the HR for living alone is 1.07 but I calculate the RR as 1.81. Conversely, 35% of survivors and 33% of deceased are current smokers, yet the HR is 1.71.

5. In the opposite direction, there are powerful influences of heart rate, despite the median being 76 in both groups.

6. The measures of blood pressure clearly show that the 474 GPs in the study used digit preference. But this in turn raises the issue of inter-observer variability when the mean number of subjects is 3 per doctor.

7. The issue of laboratory standardization was raised by Reviewer 1 but not answered.

8. The concept of ‘person-years deceased’ (Results line 6) is a curious one.

9. Table 4 should include confidence intervals.

10. It is bizarre to include data from Table 1 as a column in Table 6.

I list below the queries of the Reviewers 1 and 3 which I feel were not satisfactorily answered.

Reviewer 1
Any valid standardization of lab values?

"In table 4 the highest 5 years CVD mortality was noted in patients with good SRH versus excellent SRH, but the HR was much lower for the category "poor". Date by chance to due to much smaller sample size? By contrast the predictors of all-cause mortality in patients who died within 3 years of diabetes diagnosis are more plausible."

The authors reply that ‘In the subgroup analysis in Table 5 (new 4) of “healthy” patients without any known chronic conditions or complications (from a specified list, see Table 5 (new 4)) at diabetes diagnosis, the trend still exists in the univariate analysis and disappears in the multivariate model’,

Is this a formal test for trend or are they merely eyeballing the numbers? The same question applies on p10 line 1 with regard to Table 1.

Reviewer 3

(1) “Although DM was newly diagnosed in 1323 patients at the baseline of the study, 298 (22.5%) patients of them died during a 5-year follow-up. Even in relatively healthy subjects without any other diseases (n=696, Table 5), the mortality was 14.5% (101/696).”

While the authors have responded to the issue of treatment, they do not tackle the question of the enormously high mortality. The 5 year mortality in the conventional treatment limb in ADVANCE (mean age 66) was 9.6%, 3.5 year mortality in ACCORD (mean age 62 with known CV disease) 4%, and 10 year mortality in UKPDS (mean age 53) 18.9%.

(2) “A part of patients had cardiovascular disease (378 patients) or cancer (63 patients) at baseline of the study. Naturally, the two diseases influence baseline serum tests, BMI, heart rate, blood pressure, lifestyles and SRH. Subsequently the two diseases increase mortality of the patients, and 141 (51%) patients out of 278 patients deceased had cardiovascular disease at baseline (Table 1). In such a DM population with other severe diseases, predictors of mortality of DM patients can not correctly be drawn. Therefore, the patients with the two diseases at baseline should be excluded from the study.”

The issue of confounding, and residual confounding, is made in my point 1.

(7) “Authors/scientists should estimate statistical models based on their hypothesis and previous evidence before the beginning of analysis. Post-analytic selection of the models using AIC/BIC may be just statistically logical, but may neither be always biologically logical nor be related with authors’ hypothesis. For an example, comparison of results from Model 1 and Model 2 in Table 1 suggests that SRH are associated with yes/no of cardiovascular disease and subsequently that OR by SRH and OR by yes/no cardiovascular disease are confounded with each other. Therefore, Model 2 may not be suitable for the study objectives.”
While this reviewer feels that Model 1 is a reasonable test of their hypothesis, and Model 2 attempts to correct for confounding, the views of a statistical reviewer would be important to judge the merit of the authors’ response.

(8) “Generally, statistical power for detecting interactions (p.8) between two variables is low. The sample size (n=1323) is too small correctly to test interactive effects.”

The authors would be able to increase their numbers of events by looking at 10y mortality. This might also allow them to look at cause-specific mortality, and in a way which obviates their very crude lumping together of different disease processes in something such as ‘CVD mortality.’