Reviewer's report

Title: Statins and the risk of type 2 diabetes mellitus: cohort study using the UK Clinical Practice Research Datalink

Version: 2 Date: 5 March 2014

Reviewer: Carol Coupland

Reviewer's report:

The authors have made many changes to the paper which have improved it considerably.

They have addressed most of my previous comments very well although there are still a few points remaining.

Major Compulsory Revisions

1. “Missing data were quantified for each variable and when the nature or the extent of missingness was considered important (e.g. BMI) we conducted sensitivity analyses to explore its impact on the results.”

The authors have still not adequately explained how they accounted for missing data in the analyses. They need to clarify whether their main analyses were complete case analyses (i.e. excluding all patients with missing data for any of the confounding variables), and also specify in the Methods precisely what sensitivity analyses they carried out. This is important in particular for BMI which is likely to be a major confounder. Also by treating it as a categorical variable they may not have adequately controlled for its confounding effect.

2. Now that rates have been added to Table 4 we can see that the incidence rate of diabetes in statin users is fairly constant across all age bands, whereas there is a steep increase with age in non-statin users. This seems odd and rather implausible, and needs comment and some explanation.

3. The duration categories given in the Methods now match with the categories in the Results and Tables 4 and 5. It would have been better though to combine the 15-20 and 20-25 year categories since there are so few events in the 20-25 year category. The authors also have not clarified whether the duration of exposure variable (also called follow-up time) was treated as a time varying exposure in the analyses – this is important since otherwise results are susceptible to immortal time bias.

Minor Essential Revisions

4. “We are reluctant to present NNH because these might be taken to imply that we believe the findings were causal. Typically NNH are derived from RCTs which is not the case here.”

It is strange that the authors have made this comment about NNH when they
have already included NNH values in the manuscript although it is not an RCT. If they are keeping NNH in the paper then I think they should add them for age group, since they have given them for people with no diagnosed hypertension or cardiovascular disease. They should also add the caveat that these figures might imply causal findings. It would also be helpful to give some more explanation on how they were calculated, since various approaches can be used with survival type data and adjusted hazard ratios.

5. The authors say in the Methods that “Secondly, we restricted the analysis to the first 6 months of exposure, since a positive association over a short exposure duration could indicate a possible bias.” There is very little mention of the results of this analysis in the Results section (“The hazard ratios were higher in the first 6 months of exposure”), and very little in the Discussion, yet this seems to contradict their finding reported in the Abstract and elsewhere that “Statin use was associated with an increased risk of T2DM … which increases with longer duration of use”. Finding a larger hazard ratio in the first 6 months of use does suggest detection/ascertainment bias which might be most marked soon after starting statins but continue throughout use. Because of this I think the authors overstate their conclusions in several places (e.g. “The increased risk of T2DM should be taken into account when considering the risk-benefit balance and cost-effectiveness of statin therapy”), although they do have a more general cautious statement (“These findings should be interpreted with caution as observational studies are subject to residual confounding by indication and other biases that cannot be ruled out.”).

6. The 3rd sentence of the 4th paragraph of the Results still is unclear – “were being prescribed an association of OAD and insulin”. It would be better to change “association” to “combination”.

7. There is a hazard ratio of 170 in Table 3, which I presume should be 1.70.

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests: None.