*Response to Reviewers comments.*

Dear Dr. Goldberg,

Thank you for submitting the above manuscript to EUROPEAN UROLOGY.

We are pleased to inform you that, although your manuscript is not considered suitable for publication in its current form, we would be open to reconsider it after it has undergone major revisions. Please note that less than 10% of submitted articles are sent back to the authors with the opportunity to be reconsidered after revisions.

Your paper has been peer-reviewed by our expert reviewers and the Editorial Board.

Based on these opinions, please find below the reviewers' comments that require addressing.

If you choose to revise your paper, you should submit the modified version to the above site using your username and password and edit your existing submission.

You should take into careful consideration each point raised by the reviewers and editors. During the resubmission process, please provide a detailed explanation describing the modifications made to the manuscript. The document containing this explanation should be added as a separate item called "Revision Notes". Please also make sure the modifications are noted in the revised submission either by tracking marks or by highlighting the text

If you decide to resubmit your paper, your modified version should be received within 30 days from today's date.

Once the revised version of your manuscript has been received, it will undergo a rigorous review process.

Please note that it is possible that your revised manuscript will undergo an Editorial Statistical Review at this stage so please anticipate additional comments which will need to be addressed after this round of review. Thank you for your interest in and support to EUROPEAN UROLOGY, Your Platinum Journal.

Kind regards,

James Catto Matthew R. Cooperberg

Editor-in-Chief Associate Editor

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Comments to Author:

Reviewer #1: This manuscript assesses whether two pharmacologically-defined subgroups of statin medications, hydrophilic and hydrophobic (lipophilic) statins, show differential associations with prostate biopsy, total prostate cancer risk, and prostate cancer mortality. It leverages a large population-based cohort in Ontario defined by the local health plan, coupled with prescription drug data. The study considers men to be at risk of these outcomes if they underwent a prostate biopsy between 1994 and 2016 and did not have a diagnosis code or billing code for prostate cancer-directed treatment recorded within three months of the biopsy. The exposure is mainly defined as time-updated ever vs. never use of hydrophilic and of hydrophobic statins. With this definition, 50% of the study population is exposed to statins at some point in time (60% hydrophobic statins; these are presumably mainly simvastatin or atorvastatin). With >7000 biopsies, >5000 prostate cancer diagnoses, and 805 deaths from prostate cancer over the course of follow-up, estimates from this study are quite precise. The authors find stronger inverse associations between hydrophilic statin use and the main outcomes than for hydrophobic statins. They report a number of exploratory analyses of other medications unrelated to the main question.

The study base with its available exposure, confounder, and outcome data is well-suited to answer the clinically important main question of choosing the appropriate statin drug for cancer-related outcomes. Such knowledge would be critical for the design of randomized-controlled trials of statins for cancer endpoints, even beyond prostate cancer. Unfortunately, at least in the present form of this manuscript, the analysis adds more confusion than it resolves. How to address these self-inflicted injuries has been well described in the methodological literature, and statins have commonly served as example applications. The following comments are intended to provide the authors with some starting points for more valid analyses and interpretations. Including a pharmacoepidemiologist on the study team would be ideal.

1) A major issue with the study design is that the population at risk, start of exposure, and start of follow-up do not align. Causal interpretations are thus difficult. The main exposure, time-varying ever vs. never use of a medication, adds more "immortal" follow-up time to "ever" users (1), because ever-usage implies that people cannot switch back into the "non-user" category. Remediating this issue is at the same time easy (in term of programming) and difficult (in terms of proper definitions). If assessing time-varying use, it would be advisable to start follow-up with the start of exposure (statin use), not at the time of biopsy. Such as design has been termed "new-user design" (Wayne Ray) and is the fundamental building block of "emulated trials" in the causal inference literature (2). Alternatively, in a much simpler model, the authors could use at-baseline statin use as the exposure, eliminating inherently more complicated designs that use post-baseline exposure.

In any case, it is critical to explicitly describe these fundamental study design decisions. (Such a description is currently absent from the methods section but can be inferred to some extent from the results tables.) A graphical representation, as described by Schneeweiss et al. (Ann Intern Med 2019), is recommended. The following articles should provide some additional introductory reading: Dickerman et al., Nat Med 2019; Weberpals et al. Eur J Epidemiol 2017; Hernán et al., J Clin Epidemiol 2016.

*We agree with the reviewer’s point that it would be better to use “new-user design” where follow-up time starts with start of exposure (statin use), as it will simplify and improve the design especially when doing a head to head comparison of the two types of statins. We have done a new analysis comparing the statins head to head as described in the point below implementing the “new-user design”.*

*Hanan: Perhaps we can include a graphical representation.*

2) The study relies on indirect comparisons between users of hydrophilic statins and hydrophobic statins. However, the main question of this study is rather simple. Among people who should be prescribed a statin anyhow, would hydrophilic statins be better than hydrophobic statins? Directly implementing this question analytically leads to a simple design of higher validity than the current one. Rather than comparing non-users of statins to subgroups of users, the study should be restricted to statin users only, and the exposure would be a binary indicator of hydrophobic statin use instead of hydrophilic statins. Follow-up should start at the time of statin initiation. A restriction to a previously negative biopsy is not necessary. It mainly harms precision without improving validity.

*We agree with this point of the reviewer and performed a new analysis where we directly compared the two statins by restricting the study to only statin users and creating a binary indicator variable of exposure where “0” indicates hydrophobic statin use and “1” indicates hydrophilic statin use. Further, as mentioned in the point above, we start follow-up at time of first statin initiation, and classify the exposure based on this as in intention to treat analysis. (There 147 individuals who took both types of statins over the course of the follow-up but they were classified based on the first statin taken). We also included the other medications as covariates modeling them as binary indicator variables of exposure at baseline, where baseline is defined as use at time of statin initiation. For example, if a person used insulin before or at the time of first statin he would be considered exposed to insulin. Further, all covariates were included in the model as baseline covariates.*

*However, the restriction to previously negative biopsy is important because one of the things we are interested in is the effect of statin medications on the incidences of prostate cancer diagnosis. To be able to do this it is important to make sure 100% that the participants do not have prostate cancer or have not been diagnosed with prostate cancer, in other words, are free from prostate cancer before including them in the study. Further, we did not have their lifelong medical information, but only from the ages of 65 years and over, hence, we required that participant had a negative biopsy between the ages 65 – 66 years before including them in the study. We do understand that this makes our population a higher risk population.*

3) Confounding control with the currently used variables is rather insufficient. The analysis just described would inherently control for confounding by indication. In analytical populations not restricted to statin users, it would be important to include data on at least clinical indications of statins, such as different manifestations of cardiovascular disease and primary hyperlipidemia, and ideally also upstream causes, such as smoking and body mass index, if available, given their influence on aggressive prostate cancer. Controlling for age (at biopsy?) in five-year categories will also leave residual confounding by age, which is strongly associated with statin use and even type of statin use. Similar considerations apply for even stronger differences by calendar year. To address confounding by age and calendar year, at least one of the two should be considered for the primary time scale of survival follow-up, instead of time since biopsy. See Cologne et al., Epidemiology 2012 for some intuition.

*We agree with the reviewer that confounding by indication is indeed a valid point. We have done the above analysis, which does a head to head comparison restricting to only statin users and mostly resolves confounding by indication. Further, it is a limitation of the dataset that there is no data on different manifestations of cardiovascular disease, primary hyperlipidemia, or upstream causes, such as smoking and body mass index. And we will include this point in the limitation section.*

*To address possible residual confounding by age and calendar year, we cannot do exactly what the reviewer asked due to the limitation of the dataset which is missing exact calendar and birth dates due to data de-identification, and only includes exact number of days from an index date, which is the date of biopsy. This is noted as a limitation. However, in the new analysis where we compare the statins head to head we included as a covariate a variable which is the time from of study entry till time of statin initiation, which takes care of the possible confounding effect of duration from biopsy to statins initiation.*

4) Relying on primarily on statistical significance testing for inference about the main question is ill-advised. For example, for the most informative outcome of prostate cancer death, the hazard ratio for time-varying ever vs. never use of hydrophobic statins is 0.83 (95% CI, 0.67 to 0.98), while the hazard ratio for hydrophilic statins is 0.68 (95% CI, 0.53 to 0.87). Because the first result does not meet a Bonferroni-based p-value cut-off, it is not even mentioned in the manuscript text, and it is interpreted as support for the notion that only hydrophilic statin use would be "protective." This is a rather extreme example of misinterpretation of hypothesis tests. As the confidence intervals suggest, heterogeneity between these two estimates cannot be concluded. (And as noted above, these estimates are invalid to begin with.) Bonferroni correction is strongly discouraged in this study altogether, as the focus should be validity of estimates, not decision making for regulatory purposes, as it would be for the primary analysis in a clinical trial. It is also noted that the "negative control" analysis is not actually null but shows moderately strong associations, indicative of residual confounding. Dismissing those becomes easier if focusing on "significance thresholds." See Rothman, Gen Intern Med 2014, specifically points 5 and 6.

*We agree with the reviewers point and in the new analysis of the head to head comparison where we aim to measure associations, and infer whether they reflect causal connections, we will focus on the magnitude of these associations as the primary goal: estimation of effects is decidedly preferable to statistical testing as is mention in Rothman, Gen Intern Med 2014.*

Some other comments

5) Total prostate cancer is not a very informative endpoint and should be deemphasized.

*We agreed with the reviewer’s point.*

6) Tables from multivariable models should not report estimates for other variables, as these have no meaningful interpretation in models designed to reduce confounding for the main exposure of statin use. In contrast, event numbers and follow-up time should be reported within strata of the exposure. See Westreich and Greenland, Am J Epidemiol 2013; Rothman, Eur J Epidemiol 2017.

*We will include the event numbers and follow-up time in the tables as requested here by the reviewer but do believe covariate effect estimates would be of interest to the reader.*

7) How was "mean follow-up (SD)" calculated? It would be advisable to report median follow-up from a reverse Kaplan-Meier estimator.

*We agree with reviewer and have calculated the mean follow-up using from reverse Kaplan-Meier estimator.*

8) Causality-implying language should be revised: "effect of … on …" should read "association of … and …," and so forth.

*Completely agree with this point and have done the relevant changes in the manuscript.*

9) Inappropriate precision should be avoided. Hazard ratios should at most be reported with 2 significant digits.

*Agree with this point and have* *done the relevant changes in the manuscript.*

10) The discussion needs considerable revision. Most of the space should be spent on interpreting and discussing results in the context of the pharmacoepidemiologic literature and the limitations of the current study. In-vitro studies can briefly be mentioned as the motivation for the current study, either in the introduction or the discussion. The current "limitations" section needs fundamental revision, as most points listed are "pseudo-limitations." The main, specific potential issues with validity and generalizability of the current study are missing.

*Agree with this point and have* *done the relevant changes in the manuscript.*

Reviewer #2: The authors examined the association between different types of statins and prostate cancer risk.

1. At the least, a sensitivity analysis should be done using competing risks for all cause mortality.

*Our main interest was in associations measured through cause specific-hazards rather than absolute risks of prostate cancer, and thus we used the Cox model for the main analysis (ref: Austin, Lee & Fine 2016). However, we note in the discussion section that subdistribution hazard analysis using the Fine & Gray model would be an alternative if the interest is to determine factors that are associated with cumulative incidence of prostate cancer.*

2. It does not appear p-interactions were tested between hydrophobic and hydrophilic? To me, the most crucial outcome is PC death which is lower in both statin groups. While I realize it was only "significant" in the hydrophilic, the general direction is lower in both. Thus, unless there is a significant interaction, the overall claims of the paper that hydrophilic statins "work" but hydrophobic don't is not substantiated.

*We have performed a head to head comparison analysis of the two statins; the new analysis directly answers the question if hydrophilic statins should be used over hydrophobic, which removes the need for p-interactions.*

Reviewer #3: This is a large study of hydrophilic and hydrophobic statin use with respect to prostate cancer incidence, prostate cancer biopsy, and prostate cancer mortality using the linkage of administrative data in Ontario. Strengths of the study include the well described biologic rationale for evaluating the two classes of statins separately, the ability to capture statin use as a time-dependent and cumulative exposure, and the incorporation of a number of sensitivity analyses. A few changes or additions to the analysis and interpretation of the results would further strengthen the paper.

1) The focus of this paper seems to be on the analyses of PCa incidence, but the study lacks the information needed to overcome concerns about detection bias for this outcome. In fact, analyses of prostate biopsy are consistent with differential detection, and no data available on stage/grade at diagnosis is available. However, it is noteworthy that associations of hydrophilic statins and PCa mortality are stronger than for PCa incidence or prostate biopsy, suggesting that hydrophilic statins are associated with clinically significant disease. This could be highlighted much more in the results and discussion. In fact, re-framing the manuscript to focus on PCa mortality as the primary analysis would make a much more compelling paper.

*We agree with the reviewer and will emphasis more on PCa mortality.*

2) In the Study design, setting, and participants section, the authors list two justifications for restricting to men with a prior negative prostate biopsy. The first justification of the is: "These men are at an increased risk to develop PCa as seen in the PLCO trial, showing men with a negative biopsy having a PCa-specific mortality rate of 2.93 fold higher than men in the general population." The interpretation of the Lewicki et al study cited is a bit misleading in that: 1) the effect estimate presented is about dying from PCa, not being diagnosed with PCa; 2) the sHR was not adjusted for any other covariates, including PSA, which presumably would have strongly attenuated the association because the negative biopsy group was taken from the PLCO screening arm; and 3) the absolute difference in mortality rates in the control group and negative biopsy group are small. The second justification is: "This was used as a pre-screening tool to include a "healthier" population seen fit to undergo a biopsy." It seems that the real justification for analyses of prostate cancer incidence and biopsy, and perhaps mortality to a lesser extent, is to try to account for PSA screening prior to cohort entry/statin use, which could confound the association.

*Hanan: Could you address this comment since did not do this. One thing you could mention is that few people had PSA information in the dataset which is why it was not included and this is a limitation of the study. Our data includes people from 1964 and PSA data was only available after 2007 which significantly decreases sample size and hence was not included in the model as covariate.*

3) The authors incorporated PCa treatment into multivariable models of PCa mortality, but do not elaborate on how these variables were included. Was only treatment that occurred prior to statin use included? Including treatment that occurred after statin use could induce a selection bias if there are unmeasured factors related to both treatment and PCa mortality.

*We agree with this point. In the new analysis comparing the statins head to head the PCa treatment were included as baseline covariates hence only included if occurred prior to or at the time of statin initiation. As a side note, all covariates in the new head to head comparison analysis have been included as baseline covariates as mentioned previously under Reviewer #1 Comment 2.*

4) Is it possible for the authors to evaluate associations with overall mortality? Given that the exposure likely affects other causes of mortality, it would be helpful to be able to interpret the reduction of PCa mortality in relation to other causes of death. The authors could also consider adding Fine & Gray subdistribution HR estimates as supplementary material.

*Please see our response to similar point from Reviewer #2.*

*Hanan: I believe we did do a model with overall mortality as the outcome. Do you remember what the results was?*

5) In the Discussion on page 16, the authors state that, "…this was a retrospective population-based analysis with its associated selection bias." The authors should provide a more detailed description of how selection bias, rather than generalizability, could affect their results.

*We agreed with the reviewer completely. The issue is not selection bias but rather generalizability. We have made the relevant changes in the manuscript to reflect this.*

*Hanan:This should be corrected.* ***Should not state that there is selection bias****. Restricting to a subpopulation such as people with negative-biopsy is not selection bias. This issue is a generalizability of results to non-negative biopsy people. Generalizability limitation: all our results are valid but only for the higher risk population who have a negative biopsy. Hence, generalizing our results to those who do not have negative biopsy population is the issue. This is very different from selection bias.* ***Selection Bias*** *would be if within the higher risk population of negative-biopsy people there were only extremely sick people selected, which is not the case since everyone, sick or not, was included if they met the negative-biopsy criteria.*