## PEER REVIEW HISTORY



## ARTICLE DETAILS

| | |
|---|---|
| **TITLE (PROVISIONAL)** | Legacy effects of statins on cardiovascular and all-cause mortality - A meta-analysis |
| **AUTHORS** | Nayak, Agnish; Hayen, Andrew; Zhu, Lin; McGeechan, Kevin; Glasziou, Paul; Irwig, Les; Doust, Jenny; Gregory, Gabriel; Bell, Katy |

## VERSION 1 – REVIEW

| | |
|---|---|
| **REVIEWER** | Jose Gutierrez<br>Columbia University, USA |
| **REVIEW RETURNED** | 27-Dec-2017 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors embarked on a project to explore whether statins are associated with a post-trial legacy effect. The question is pertinent, and the result do partially address the question but there are many uncertainties that should dampen the enthusiasm of the conclusion about the use of statins early on to prevent vascular disease. There exist current guidelines that provide guidance to clinicians on when to start statins, and although not free of controversy, these guidelines do include multiple other sources other than trials. Nonetheless, I believe this work is important and I have some comments to the authors.<br>-Is there a reason to limit the study to cohorts > 1000 participants?<br>-Data search strategy: Authors should report the search strategy used for identifying additional trials in Medline and embasse.<br>-was there any correlation between point estimates and difference in statins ascertainment percentage within trial and post-trial?<br>-How do author explain the consistently discrepant statin effects of 4s trial in within- and post-trial point estimates?<br>-Although figure 1 is educational, it may lead to confusion. Some may believe it is an actual result rather than a simulation. Perhaps it would fit better in the supplementary data.<br>-Is there any relationship between the proportion of women included in each study and the magnitude of the benefit of statins in the post-trial period? Those with higher proportions of men could potentially have higher benefits as benefits of statins seem attenuated in women, at least for secondary prevention. Interestingly, WOSCOPS trial had the largest weight in the primary prevention estimates, and it had no women. Allhat had the largest proportion of women, and the lowest risk reduction. A sex-stratified meta-analysis would be prudent, if the studies reported sex-based estimates. A sex-effect is not excluded, and thus should be a highlight of the discussion.<br>-A source of uncertainty about what drives the theoretical legacy effects of statins may relate to differences in follow-period by trial and age at the time of randomization. Older cohorts with longer follow up would possibly have an attenuated risk reduction compare with younger cohorts. A scattered plot of point estimates using age and follow-up time may be helpful to assess this possibility visually. |

| | |
|---|---|
| | -It is interesting that ASCOT has the lowest point estimate for overall mortality in the primary group. ASCOT included people with > 3 vascular risk, which would have a higher risk of events as predicted by Framingham score, probably >10% in 10 years, which is considered a coronary equivalent. So it raises the question of the artificial separation or inclusion of ASCOT as primary prevention trial. |

| | |
|---|---|
| **REVIEWER** | Irtiza Hasan<br>University of Texas MD Anderson Cancer Center, Houston, Texas, USA |
| **REVIEW RETURNED** | 03-Jan-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | The manuscript is well written and the level is appropriate to our readership. The subject is very important and the findings have clinical significance.<br><br>In my view, this manuscript ticks almost all the boxes we normally have in mind for a meta-analysis paper and I am recommending few corrections to improve the manuscript. Please find the details below.<br><br>The introduction needs to be improved particularly with more background information.<br><br>Methods need to be more detailed, especially the search strategy (box 1), validity assessment & measure of consistency between the studies included. Also, it would be nice to mention whether authors who checked the titles and abstracts, as well as the full-text articles, were blinded and what was their convex alpha for agreement. Authors need to mention whether & what tools they used to assess the study quality.<br><br>In results section (on page 11), for all-cause mortality, the second line and the last line related to WOSCOPS study contradicts; in the second line, it says the WOSCOPS showed less benefit in post-trial period whereas, in the last line it says, the trial showed that there was a significant reduction in all-cause mortality. The authors didn't mention whether or how they looked at the publication bias. Majority of the studied had a post-trial follow-up period of fewer than 6 years. Only 2 of the trials which showed significant benefit had a follow up of 15.1 & 8.3 years. Does the short duration of post-trial follow up have any role in the benefit detection? The page numbers mentioned for most of the section/topic on the PRISMA 2009 checklist are not correct.<br><br>The authors could strengthen the paper by adding a conclusion. |

| | |
|---|---|
| **REVIEWER** | Apostolos Tsapas<br>Aristotle University Thessaloniki |
| **REVIEW RETURNED** | 01-Feb-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | In their systematic review and meta-analysis, Nayak et al evaluated the legacy effect of statins on CVD and all-cause mortality. They only considered randomized studies; they performed sensitivity analysis restricting the model to data from primary prevention studies. The coverage of literature was accurate and up-to-date. The overall take-home message is clear and well substantiated.<br><br>Minor comment |

| | |
|---|---|
| | Please double check eFigure for consistency: Initial search results do not add up to 2413 |

| | |
|---|---|
| **REVIEWER** | Jeanine Roeters van Lennep<br>Department Internal Medicine, Erasmus MC, Rotterdam, The Netherlands |
| **REVIEW RETURNED** | 11-Feb-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | This meta-analysis assessed the evidence of a “legacy” effect of statin-therapy on cardiovascular disease and all-cause mortality outcome in primary and secondary prevention randomized controlled trials which compared statins to placebo. The main finding was that in that compared to the secondary prevention RCTs, in the primary prevention RCTs some evidence of legacy effects on all-cause mortality and was found. Moreover the pooled post-trial hazard ratio for the primary prevention trials showed possible post-trial legacy effects on CVD mortality and all-cause mortality. The conclusion of the authors was that this indicates that early treatment of atherosclerosis is likely to be beneficial.<br><br>**General comments:**<br><br>-The author focus solely on the LDL-C lowering effect of statins. It would be interesting to discuss whether pleiotropic effects of statins (anti-inflammation, anti-coagulation effects) could play a role in the legacy effect.<br>- If no legacy effect is found in secondary prevention trials, this would imply that discontinuation of statin therapy will cause harm on CVD and all-cause mortality. The authors do not discuss this aspect at all. It would be adding value if they did discuss view point as well<br><br>- The results of legacy effect are largely driven by the primary prevention WOSCOPS trial which consisted of only men and had the longest follow-up. Maybe the effect is stronger in primary prevention trials but maybe it is the follow-up time of maybe the effect is only present in men?<br>It would be interesting if more exploratory sub-analyses on legacy effects could be performed to study for instance effect of length of follow-up during trial as well as post-trial follow-up, statin-intensity, number of patients using statins post-trial or ratio statin users within vs posttrial in patients initially allocated to statins and placebo and percentage women in trial<br><br>**Specific comments**<br><br>-Was there a minimum time of follow-up defined of follow-up time within and after trials?<br>-For clarity it would be helpful to also show survival and hazard ratio curves in if a legacy effect does exist in FIgure 1a/b.<br>-It would be helpful to arrange the trials in primary and secondary prevention trials in figure 2,3 and 4. |

| | |
|---|---|
| **REVIEWER** | Robert M West<br>University of Leeds, UK |
| **REVIEW RETURNED** | 07-Mar-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | There are some major issues with this work so that I recommend |

| | |
|---|---|
| | rejection. The work is well done but the major issues unfortunately prevent any other recommendation.<br><br>I note that the review protocol was not registered.<br><br>Articles were selected on the basis that they reported RCTs with at leat 1000 participants. I assume that this is to ensure research quality data. The legacy effect though is studied in the post-trial period and does not fully benefit from randomisation. Would it have been more important to select on the quality (and duration) of the post-trial follow up? The criteria used have the advantage of limiting to a manageable number of publications to be considered but appear to me the wrong criteria.<br><br>The exposure to statins post trial is not known. This makes it extremely difficult to assess the findings. What value do the findings have?<br><br>The simulations of idealised scenarios illustrates the effect sought but could be misleading - the situation studied is not likely to be ideal.<br><br>The authors have a complex argument for the importance of legacy effects and why they are interesting. I query if this argument is sufficiently convincing to merit publication. In very practical terms statins, I understand, are prescribed 'for life' rather than for a limited period. In that scenario , there is little practical interest in legacy effects.<br><br>Eight trials are restricted to a subset of only three for subset analysis, so conclusions must be limited.<br><br>The wording in the conclusion assumes that if p>0.05 then there is no effect. This is not an appropriate interpretation, and rewording would be essential. This is out of character with the thoughtful presentation otherwise.<br><br>In conclusion, the exposure post trial is unknown, the selection of articles is not the most appropriate, and results of a limited number of articles over-interpreted. |

| | |
|---|---|
| **REVIEWER** | Benn Sartorius<br>University of KwaZulu-Natal, South Africa |
| **REVIEW RETURNED** | 18-Mar-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | The authors have conducted a systematic review of the legacy effects of statins on all cause and cause-specific mortality based on clinical trials. They have adhered to PRISMA guidelines and included a checklist in this regard in the appendix with index to relevant sections of the manuscript speaking to these items.<br><br>Comments/queries:<br><br>Protocol and Registration “The review protocol was not registered” Why? This is unusual as most systematic review protocols are registered with Cochrane or PROSPERO?<br><br>Search strategy: How many hits found in Medline/Embase that were not found in Scopus? Did you also check clinical trial registries e.g. Cochrane Central Register of Controlled Trials or ClinicalTrials.gov |

| | |
|---|---|
| | etc?<br><br>Selection “We performed a systematic search and meta-analysis of all reports on follow-up after randomized, placebo-controlled studies of adults (age >18 years) of statins with ≥1000 participants.” Why were studies with <1000 participants excluded? Meta-analyses may include data from smaller studies which, individually, do not have power to detect a modest intervention effect. Many argue for including small studies in meta-analyses as evidence synthesis is best informed by all reasonably unbiased evidence.<br><br>Study selection and data abstraction: “Two authors (AN and KB) checked the titles and abstracts of all citations identified through the database searches and forward citation search.” Not explicitly mentioned but I presume these two authors worked independently when performing this screen?<br><br>Study selection and data abstraction: Were agreement statistics calculated for the two reviewers e.g. Kappa? I don’t see this mentioned nor presented.<br><br>Were the quality of the individual studies included assessed using accepted checklists e.g CONSORT (Consolidated Standards of Reporting Trials) statement?<br><br>Was any sensitivity analysis performed using duration of follow-up as a co-variate and its impact on the outcomes measured?<br><br>Simulation: were bootstrapped 95% uncertainty margins also estimated? These should be included in figures 1/2. |

| | |
|---|---|
| **REVIEWER** | Fazeel Siddiqui<br>SIU School of Medicine |
| **REVIEW RETURNED** | 04-Apr-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | The article is well written and addresses an important question about the legacy effects of statin use on the patients who were initially enrolled in statin trials. The meta-analysis was unable to adjust for follow up duration which is critical and major limitation. This article also did not address (mentioned briefly in the limitation ) whether there was any other factors post trial (due to lack of randomization) which could affect the cardiovascular or all cause mortality besides statins. Although the results of this meta-analysis are interesting, and worth reporting, the discussion section should be shortened and focus on how the severe limitation with the analysis (aggregate date, lack of adjustment for follow up time period and other factors that could affect CVD and all cause mortality) prevent us from generalization of these findings. |

**VERSION 1 – AUTHOR RESPONSE**

**REVIEWER COMMENTS:**

**Jose Gutierrez (Reviewer 1):** The authors embarked on a project to explore whether statins are associated with a post-trial legacy effect. The question is pertinent, and the result do partially address the question but there are many uncertainties that should dampen the enthusiasm of the conclusion about the use of statins early on to prevent vascular disease. There exist current guidelines that provide guidance to clinicians on when to start statins, and although not free of controversy, these guidelines do include multiple other sources other than trials. Nonetheless, I believe this work is important and I have some comments to the authors.

Response: Thank you for your supportive comments.

Change: Nil

-Is there a reason to limit the study to cohorts > 1000 participants?

Response: Our included studies were all follow-up reports of trials included in the Cholesterol Treatment Trialists' Collaboration, and one of their inclusion criteria is >1000 participants in the trial. We have added this rationale to the Methods.

Change: Page 8. Lines 148-152: "We chose to limit our studies to those with ≥1000 participants in the original trial for consistency with the Cholesterol Treatment Trialists' Collaboration. These large trials were designed to assess effects on mortality within the trial period, and their follow-up reports are the most appropriate studies to address post-trial effects on mortality."

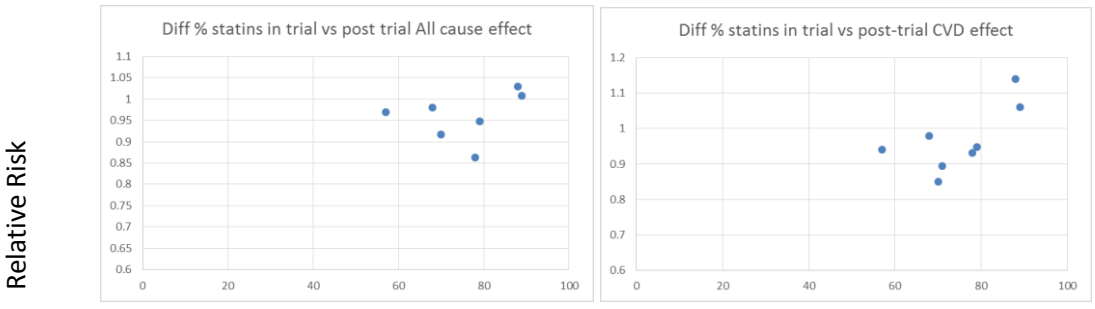
-Data search strategy: Authors should report the search strategy used for identifying additional trials in Medline and embasse.

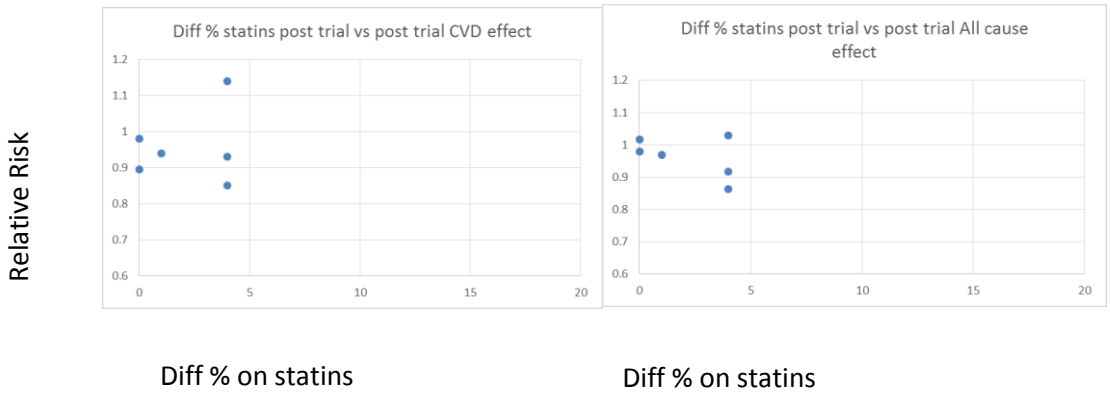
Response: Our search terms for Medline are provided in Box 1. We have not provided additional search terms for Embase, but are happy to do so if the editor requires this.

Change: Nil.

-was there any correlation between point estimates and difference in statins ascertainment percentage within trial and post-trial?

Response: As you can see from the scatterplots on the next page, there are no clear correlations observable. However the limited number of data points make it difficult to draw meaningful conclusions.






Change: Nil

-How do author explain the consistently discrepant statin effects of 4s trial in within- and post-trial point estimates?

Response: We assume the reviewer is referring to the findings of a large beneficial effect within trial and no evidence of effect post trial. The results for 4S follow are actually consistent with all of the studies whereby the within trial estimates are much larger than the post-trial estimates, a result we emphasise in the Discussion. Although the point estimate post trial is >1, we do not interpret this as suggesting a potentially harmful legacy effect as the confidence intervals clearly overlap 1 (null effect). In addition, as mentioned in the Discussion, the post-trial comparison is no longer randomised and the group originally allocated to statin will actually be at higher risk independent of any legacy effect.

Change: Nil

-Although figure 1 is educational, it may lead to confusion. Some may believe it is an actual result rather than a simulation. Perhaps it would fit better in the supplementary data.

Response: We agree and have moved this to an online only supplementary file.

Change: Figure 1 moved to online only supplementary file and Figures renamed accordingly.

-Is there any relationship between the proportion of women included in each study and the magnitude of the benefit of statins in the post-trial period? Those with higher proportions of men could potentially have higher benefits as benefits of statins seem attenuated in women, at least for secondary prevention. Interestingly, WOSCOPS trial had the largest weight in the primary prevention estimates, and it had no women. Allhat had the largest proportion of women, and the lowest risk reduction. A sex-stratified meta-analysis would be prudent, if the studies reported sex-based estimates. A sex-effect is not excluded, and thus should be a highlight of the discussion.

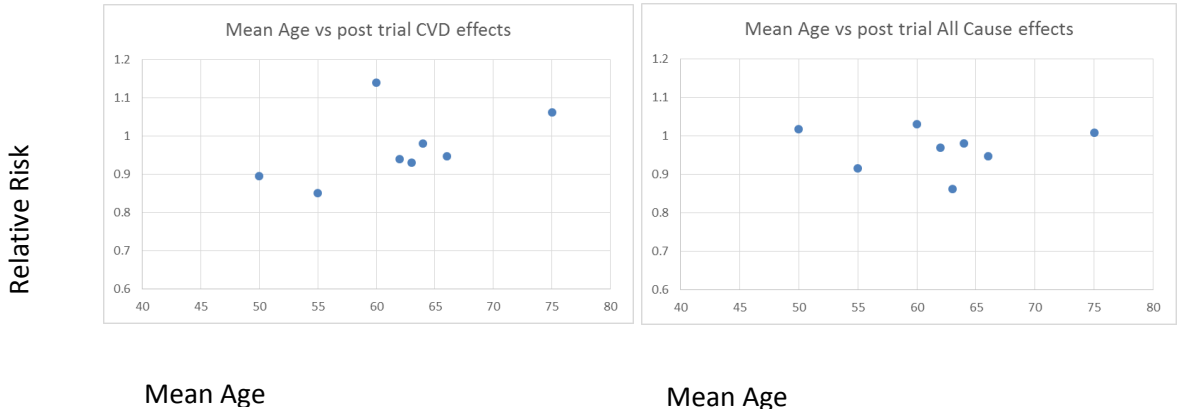
Response: We agree that this would be interesting to explore, but unfortunately we do not have sufficient data to undertake sex-stratified analysis. This is one issue of trial level meta-analysis and we have now included some discussion on this in the paper.

Change: Page 16, lines 318-322: “For example although we found evidence of possible legacy effects in primary care, these are largely driven by WOSCOPs which was undertaken in all male participants. If there are sex-specific effects for legacy effects, it may be the fact that all participants in WOSCOPS were male, and not that they had no history of CVD, that is the more important determinant.”

-A source of uncertainty about what drives the theoretical legacy effects of statins may relate to differences in follow-period by trial and age at the time of randomization. Older cohorts with longer

follow up would possibly have an attenuated risk reduction compare with younger cohorts. A scattered plot of point estimates using age and follow-up time may be helpful to assess this possibility visually.

Response: We purposively ordered the trials (within primary and secondary prevention strata) by length of follow-up in all the Figures, to allow visualisation of effect of longer follow-up. As for age, from the scatterplots below we can see no clear correlations with legacy effects. As before, the limited number of data points make it difficult to draw meaningful conclusions about importance of either of follow-up time or age.




Change: We have included the following note for Figures 1, 2 and 3: “Note: Within primary and secondary prevention subgroups, studies are ordered by duration of follow-up.”

-It is interesting that ASCOT has the lowest point estimate for overall mortality in the primary group. ASCOT included people with > 3 vascular risk, which would have a higher risk of events as predicted by Framingham score, probably >10% in 10 years, which is considered a coronary equivalent. So it raises the question of the artificial separation or inclusion of ASCOT as primary prevention trial.
Response: We agree that this is interesting. But in fact as we note in the Discussion all three primary prevention trials were of people who are likely to have risk estimates above current treatment thresholds. We have included some event rates from the placebo groups in the trials to show that these are high risk people.

Change: Page 17, lines 357-361: “For example, the proportion of people who had died of cardiovascular disease by the end of the trial in the placebo group after 3.3 years in ASCOT, 4.8 years in ALLHAT and 4.9 years in WOSCOPS was 3%, 11% and 2% respectively.”

**Irtiza Hasan (Reviewer 2):** The manuscript is well written and the level is appropriate to our readership. The subject is very important and the findings have clinical significance. In my view, this manuscript ticks almost all the boxes we normally have in mind for a meta-analysis paper and I am recommending few corrections to improve the manuscript. Please find the details below.

Response: Thank you for your supportive comments

Change: Nil

The introduction needs to be improved particularly with more background information.

Response: We have included some more background information

Change: Please see specific tracked changes in Background section

Methods need to be more detailed, especially the search strategy (box 1), validity assessment & measure of consistency between the studies included. Also, it would be nice to mention whether authors who checked the titles and abstracts, as well as the full-text articles, were blinded and what was their convex alpha for agreement. Authors need to mention whether & what tools they used to assess the study quality.

Response: We have relabelled box 1 to indicate these are the search terms used in the Medline search, and are happy to include other information here that the editor feels is needed.

We did not formally assess quality of the reports. Our included studies were all from the Cholesterol Treatment Trialists' Collaboration. ROB assessments of the RCTs have been done previously by others, and we now include a recent reference that includes ROB assessment for all of the RCTs that resulted in the follow-up studies included in our meta-analysis.

We confirm that all steps of the review were initially done blinded to the other reviewer's results, and followed by unblinded discussion to resolve any differences. We did not calculate formal measures of agreement to describe the extent to which assessments by multiple authors were the same. The Cochrane handbook recommends against doing so as disagreement about the eligibility of a large, well conducted, study will have more substantial implications for the review than disagreement about a small study with risks of bias. Instead, the handbook recommends exploring reasons for any disagreement which is the approach we took through discussion.

Change:

Title of Box 1 changed to "Search terms for Medline Search"

Page 15, lines 305-308: "We did not assess risk of bias for the included studies, but this has been assessed by others for the original trial reports, including very recently(40), and the included studies were generally found to be high quality."

Page 10, lines 175-178: "We did not calculate formal measures of agreement to describe agreement between reviewers. The Cochrane Collaboration recommends against doing this, and instead recommends exploring reasons for any disagreement early on in the review process(17), which we did through discussion."

In results section (on page 11), for all-cause mortality, the second line and the last line related to WOSCOPS study contradicts; in the second line, it says the WOSCOPS showed less benefit in post-trial period whereas, in the last line it says, the trial showed that there was a significant reduction in all-cause mortality.

Response: The results are not contradictory, although the benefit post-trial was significant in WOSCOPS, it was smaller than the benefit observed within the trial period. We have made some changes to the text to make this clearer.

Change: Page 12, line 236: "showed less benefit in the post-trial period than the trial period."

Page 12-13, line2 242-243: "showed less benefit in the post-trial period than the trial period."

The authors didn't mention whether or how they looked at the publication bias.

Response: We did not assess publication bias. We mention this as a limitation in the Discussion

Change: Page 15, lines 304-305: "However we did not assess for publication bias and it is possible that unpublished follow-up reports may exist that we are unaware of."

Majority of the studied had a post-trial follow-up period of fewer than 6 years. Only 2 of the trials which showed significant benefit had a follow up of 15.1 & 8.3 years. Does the short duration of post-trial follow up have any role in the benefit detection?

Response: We purposively ordered the trials (within primary and secondary prevention subgroups) by length of follow-up in all the Figures, to allow visualisation of effect of longer follow-up. Because of the limited number of data points it is difficult to draw meaningful conclusions about this however.

Change: We have included the following note for Figures 1-3: “Note: Within primary and secondary prevention subgroups, studies are ordered by duration of follow-up.”

The page numbers mentioned for most of the section/topic on the PRISMA 2009 checklist are not correct.

Response: Thank you for alerting us to this.

Change: We have corrected the PRISMA checklist according to the revised manuscript.

The authors could strengthen the paper by adding a conclusion.

Response: We have added the subtitle Conclusion to the final paragraph.

Change: We have added the subtitle Conclusion to the final paragraph.

**Apostolos Tsapas (Reviewer 3):** In their systematic review and meta-analysis, Nayak et al evaluated the legacy effect of statins on CVD and all-cause mortality. They only considered randomized studies; they performed sensitivity analysis restricting the model to data from primary prevention studies. The coverage of literature was accurate and up-to-date. The overall take-home message is clear and well substantiated.

Response: Thank you for your supportive comments

Change: Nil

Minor comment
Please double check eFigure for consistency: Initial search results do not add up to 2413
Response: Thank you for alerting us to this error.

Change: We have corrected the numbers in the Efigure (now Efigure 2).

**Jeanine Roeters van Lennep (Reviewer 4):**

The author focus solely on the LDL-C lowering effect of statins. It would be interesting to discuss whether pleiotropic effects of statins (anti-inflammation, anti-coagulation effects) could play a role in the legacy effect.

Response: Thank you, we have added this point to the Introduction.

Change: Page 6, lines 105-106: “As with the direct effects of statins, these legacy effects may be pleiotropic, and act through anti-inflammation, anti-coagulation and or lipid lowering.”

- If no legacy effect is found in secondary prevention trials, this would imply that discontinuation of statin therapy will cause harm on CVD and all-cause mortality. The authors do not discuss this aspect at all. It would be adding value if they did discuss view point as well

Response: Thank you, we have added this point to the Discussion.

Change: Page 14, lines 287-289: “Considering these subgroups separately, we found no evidence of legacy effects following secondary prevention trials, suggesting the importance of long term /life-long prevention in these patients.”

- The results of legacy effect are largely driven by the primary prevention WOSCOPS trial which consisted of only men and had the longest follow-up. Maybe the effect is stronger in primary prevention trials but maybe it is the follow-up time of maybe the effect is only present in men?
It would be interesting if more exploratory sub-analyses on legacy effects could be performed to study for instance effect of length of follow-up during trial as well as post-trial follow-up, statin-intensity, number of patients using statins post-trial or ratio statin users within vs posttrial in patients initially allocated to statins and placebo and percentage women in trial

Response: We agree that these are all interesting hypotheses to explore, but unfortunately our data are limited to 8 studies and we think that performing more sub-analyses (even if we label these as exploratory) may be undesirable. We have however incorporated some of these points in text added to the Discussion.

Change: Page 15-16, lines 314-325: "The main limitation of our report is that because our findings are based on aggregate data, we are unable to assess the effects of whether or not an individual was treated with statins during the post-trial period, and for how long, as well as their cardiovascular risk factor levels and other confounders. For example although we found evidence of possible legacy effects in primary care, these are largely driven by WOSCOPs which was undertaken in all male participants. If there are sex-specific effects for legacy effects, it may be the fact that all participants in WOSCOPS were male, and not that they had no history of CVD, that is the more important determinant. Similarly, participants in WOSCOPS had the lowest percentage taking statins in the post trial period out of all the studies where this was measured (39% of active and 35% of placebo participants were taking statins at 5 years post-trial). This comparative absence of direct statin treatment effects in the post trial period may be the more important determinant."

-Was there a minimum time of follow-up defined of follow-up time within and after trials?

Response: No we didn't limit studies by follow up time.

Change: Nil.

-For clarity it would be helpful to also show survival and hazard ratio curves in if a legacy effect does exist in FIgure 1a/b.

Response: We now include graphs using simulated data where a legacy effect exists. For simplicity, we have set the legacy effect as half that of the direct (within trial) effect and assumed that participants do not take statins post trial. Note that these are survival and hazard curves (not hazard ratio curves).

Change pages 7-8, lines 123-136: "To illustrate this point we generated data to simulate the situation where there was, and was not, a legacy effect (we simulated two scenarios where an intervention has effects during the trial period, and (i) has an effect after the trial (legacy effect) or (ii) has no effect after the trial (no legacy effect). In the survival curves of both scenarios the apparent legacy effect is exaggerated because the cumulative incidence includes the direct effects during the initial trial period (eFigure 1A and 1C). If hazard curves are constructed instead, the direct effects during the initial trial period are not included in the instantaneous hazard of the post trial periods, allowing an unbiased estimation of the legacy effect (eFigure 1B and ID; details of the methods for the simulation are provided in the Appendix). Although survival curves like eFigure 1A and IC demonstrate that the direct effects of the intervention (observed during the trial period) are still apparent many years later, they do not provide evidence of legacy effects after the intervention has ceased. From the hazard curves in eFigure 1B and 1D it is clear that to estimate legacy effects, we should instead focus on outcomes observed during the post-trial period."

-It would be helpful to arrange the trials in primary and secondary prevention trials in figure 2,3 and 4.

Response: The trials are arranged as primary prevention first and then secondary prevention, with shortest to longest post trial followup within each of these. This ordering is consistent across Figures 2-4.

Change: Nil.

**Robert M West (Reviewer 5):**

Articles were selected on the basis that they reported RCTs with at least 1000 participants. I assume that this is to ensure research quality data. The legacy effect though is studied in the post-trial period and does not fully benefit from randomisation. Would it have been more important to select on the quality (and duration) of the post-trial follow up? The criteria used have the advantage of limiting to a manageable number of publications to be considered but appear to me the wrong criteria.

Response: As the legacy effect relates to the difference in treatment received within the trial period, we focussed on high quality, large placebo controlled RCTs of statins for our inclusion criteria (using the inclusion criteria of the LLTTC). We then included all follow-up reports of these key trials. It seems likely that legacy effects would usually be smaller than direct effects observed in the trial. If inclusion criteria were based on the follow up study rather than the original RCT, we might have ended up with studies reporting on long term outcomes after low quality trials and/or small studies with limited power to detect either direct or legacy effects. This seems to be the wrong approach to us. We have provided justification of our approach in the Methods section.

Change: Page 8, lines 147-148: “As the legacy effect relates to the difference in treatment received within the trial period, we focused our analysis on follow up reports of high quality, large RCTs.”

The exposure to statins post trial is not known. This makes it extremely difficult to assess the findings. What value do the findings have?

Response: We have included exposure to statins post trial where this was reported (See Figure 1). For example 39% of active and 35% of placebo participants were taking statins at 5 years post-trial in WOSCOPS, whereas 67% of active and 63% of placebo participants were taking statins at 2.2 years post-trial in ASCOT-LLA. We have now included some discussion on this.

Change: Page 16, lines 322-325 :“Similarly, participants in WOSCOPS had the lowest percentage taking statins in the post trial period out of all the studies where this was measured (39% of active and 35% of placebo participants were taking statins at 5 years post-trial). The comparative absence of statin treatment effects in the post trial period may be the more important determinant.”

The simulations of idealised scenarios illustrates the effect sought but could be misleading - the situation studied is not likely to be ideal.

Response: We have moved the simulations to the online only material in case it is misleading. The point of the simulation was not to be realistic, but to illustrate why the post-trial period should be the focus of analysis, rather than from the point of randomisation which has been the case in most publications (both primary studies and reviews) to date.

Change: We have moved the simulations to the online only material

The authors have a complex argument for the importance of legacy effects and why they are interesting. I query if this argument is sufficiently convincing to merit publication. In very practical terms statins, I understand, are prescribed 'for life' rather than for a limited period. In that scenario, there is little practical interest in legacy effects.

Response: We see the value of legacy effects as providing indirect evidence on two potential benefits of statins:

1. Starting statins at an earlier age (assuming that they are prescribed for life once started). As we argue in the Discussion, a trial to test this hypothesis directly is unlikely, and so post-trial

follow up studies may be the best evidence we have on this. We argue that this is of very large practical interest.

2. Longterm/lifelong treatment rather than limited period of treatment. As we indicate above, the possible legacy effects observed in WOSCOPS may be primarily because of the small number of men taking statins post-trial in this study. In the other studies where majority of participants took statins post-trial, any potential legacy benefit to patients allocated to statins in the trial, may have been negated by a greater degree of immediate protection from post-trial uptake of treatment in the patients allocated to placebo.

Change:

As well as the changes noting relative low levels of statin use post-trial in WOSCOPS as outlined above, we have added the words long term to this sentence on page 17, lines 361-363: “Legacy effects in these settings serve to emphasise the benefits of starting **long term** primary prevention treatment early rather than later among people at high short term risk.

Eight trials are restricted to a subset of only three for subset analysis, so conclusions must be limited.

Response: We agree and have emphasised that the subgroup analysis only suggests possible legacy effects.

Change: We have ensured that we use the word “possible” or “potential” to describe the results for the subgroup analysis.

The wording in the conclusion assumes that if p>0.05 then there is no effect. This is not an appropriate interpretation, and rewording would be essential. This is out of character with the thoughtful presentation otherwise.

Response: We agree that no evidence of an effect should not be interpreted as evidence of no effect, and have reviewed the paper throughout to make sure that we do not imply this.

Change: We have checked wording throughout the paper to ensure that we do not say there is evidence of no legacy effect.

In conclusion, the exposure post trial is unknown, the selection of articles is not the most appropriate, and results of a limited number of articles over-interpreted.

Response: We have addressed these issues above.

Change: Nil further.

**Benn Sartorius (Reviewer 6):**

Protocol and Registration “The review protocol was not registered” Why? This is unusual as most systematic review protocols are registered with Cochrane or PROSPERO?

Response: As our study focuses on post-trial cohorts which are no longer randomised comparisons, Cochrane was not appropriate. We were not aware of PROSPERO at the time this project commenced (at the end of 2015). We tried to register with PROSPERO in 2017 but were not able to as all steps of the review were then complete.

Change: Nil.

Search strategy: How many hits found in Medline/Embase that were not found in Scopus? Did you

also check clinical trial registries e.g. Cochrane Central Register of Controlled Trials or ClinicalTrials.gov etc?

Response: There were 1044 titles and abstracts retrieved through the forward citation search in Scopus that were not identified by the Medline/Embase database searches. The focus of our analysis was on long term follow up after published trials and so we didn't search clinical trial registries.

Change: Nil.

Selection "We performed a systematic search and meta-analysis of all reports on follow-up after randomized, placebo-controlled studies of adults (age >18 years) of statins with ≥1000 participants." Why were studies with <1000 participants excluded? Meta-analyses may include data from smaller studies which, individually, do not have power to detect a modest intervention effect. Many argue for including small studies in meta-analyses as evidence synthesis is best informed by all reasonably unbiased evidence.

Response: Our included studies were all follow-up reports of trials included in the Cholesterol Treatment Trialists' Collaboration, and one of their inclusion criteria is >1000 participants in the trial. We agree that in general it is better to include all studies no matter what size. However for this particular study question, we were interested in effects on mortality outcomes occurring post-trial in trials with sufficient power to evaluate these outcomes within trial. Thus we limited our selection to follow up reports of large RCTs.

Change: Page 8. Lines 148-152: "We chose to limit our studies to those with ≥1000 participants in the original trial for consistency with the Cholesterol Treatment Trialists' Collaboration. These large trials were designed to assess effects on mortality within the trial period, and their follow-up reports are the most appropriate studies to address post-trial effects on mortality."

Study selection and data abstraction: "Two authors (AN and KB) checked the titles and abstracts of all citations identified through the database searches and forward citation search." Not explicitly mentioned but I presume these two authors worked independently when performing this screen?

Response: We confirm that all steps of the review were initially done blinded to the other reviewer's results, and then unblinded discussion to resolve any differences. We have added the word "independently" to the description of this step of the review.

Change: Page 9, line 170: "Two authors (AN and KB) independently checked the titles and abstract"

Study selection and data abstraction: Were agreement statistics calculated for the two reviewers e.g. Kappa? I don't see this mentioned nor presented.

Response: We did not calculate formal measures of agreement to describe the extent to which assessments by multiple authors were the same. The Cochrane handbook recommends against doing so as disagreement about the eligibility of a large, well conducted, study will have more substantial implications for the review than disagreement about a small study with risks of bias. Instead, the handbook recommends exploring reasons for any disagreement.

Change: Page 9, lines 175-178: "We did not calculate formal measures of agreement to describe agreement between reviewers. The Cochrane Collaboration recommends against doing this, and instead recommends exploring reasons for any disagreement early on in the review process (17), which we did through discussion."

Were the quality of the individual studies included assessed using accepted checklists e.g CONSORT (Consolidated Standards of Reporting Trials) statement?

Response: We did not formally assess quality of the reports. Our included studies were all from the Cholesterol Treatment Trialists' Collaboration. ROB assessments of the RCTs have been done previously by others, and we now include a recent reference that includes ROB assessment for all of the RCTs that resulted in the follow-up studies included in our meta-analysis.

Change: Page 15, lines 305-308: "We did not assess risk of bias for the included studies, but this has been assessed by others for the original trial reports, including very recently(40), and the included studies were generally found to be high quality."

Was any sensitivity analysis performed using duration of follow-up as a co-variate and its impact on the outcomes measured?

Response: No. We did not have sufficient data to assess this. We did however order the trials (within primary and secondary prevention subgroups) by length of follow-up in all the Figures, to allow visualisation of effect of longer follow-up. We have now provided a note on this in Figures 2 and 3.

Change: We have included the following note for Figures 1-3: "Note: Within primary and secondary prevention subgroups, studies are ordered by duration of follow-up.

Simulation: were bootstrapped 95% uncertainty margins also estimated? These should be included in figures 1/2.

Response: The purpose of the simulations was to illustrate why the post-trial period should be the focus of analysis, rather than from the point of randomisation which has been the case in most publications (both primary studies and reviews) to date. We have chosen not to present 95% distribution of survival and hazard estimates as we do not think that this will add clarity to our point about the time period that analysis should focus on.

Change: Nil

**Fazeel Siddiqui (Reviewer 7):** The article is well written and addresses an important question about the legacy effects of statin use on the patients who were initially enrolled in statin trials.

Response: Thank you for your supportive comments

Change: Nil

The meta-analysis was unable to adjust for follow up duration which is critical and major limitation.

Response: We agree that we did not have sufficient data to assess effects of follow up duration. We did however order the trials (within primary and secondary prevention subgroups) by length of follow-up in all the Figures, to allow visualisation of effect of longer follow-up. We have now provided a note on this in Figures 1-3.

Change: We have included the following note for Figures 1-3: "Note: Within primary and secondary prevention subgroups, studies are ordered by duration of follow-up.

This article also did not address (mentioned briefly in the limitation ) whether there was any other factors post trial (due to lack of randomization) which could affect the cardiovascular or all cause mortality besides statins.

Response: We agree that we were not able to assess potential confounders in this trial level meta-analysis and have expanded the Discussion on this limitation.

Change: Page 15-16, lines: 314-322: "The main limitation of our report is that because our findings are based on aggregate data, we are unable to assess the effects of whether or not an individual was treated with statins during the post-trial period, and for how long, as well as their cardiovascular risk factor levels and other confounders. For example although we found evidence of possible legacy effects in primary care, these are largely driven by WOSCOPs which was undertaken in all male

participants. If there are sex-specific effects for legacy effects, it may be the fact that all participants in WOSCOPS were male, and not that they had no history of CVD, that is the more important determinant."

Although the results of this meta-analysis are interesting, and worth reporting, the discussion section should be shortened and focus on how the severe limitation with the analysis (aggregate date, lack of adjustment for follow up time period and other factors that could affect CVD and all cause mortality) prevent us from generalization of these findings.

Response: We have shortened the Discussion by moving some text to the Introduction. And emphasised the key limitation as above.

Change: See tracked changes to see text deleted from Discussion.

**VERSION 2 – REVIEW**

| | |
|---|---|
| **REVIEWER** | Robert M West<br>University of Leeds |
| **REVIEW RETURNED** | 04-Jun-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | Revision is good, responses to first-round queries satisfactory |

| | |
|---|---|
| **REVIEWER** | Benn Sartorius<br>University of KwaZulu-Natal, South Africa |
| **REVIEW RETURNED** | 08-Jun-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | Thanks i'm satisfied with the response to my previous comments and those of the other reviewers. |

| | |
|---|---|
| **REVIEWER** | Fazeel Siddiqui<br>Southern Illinois University Medical School, USA |
| **REVIEW RETURNED** | 04-Jun-2018 |

| | |
|---|---|
| **GENERAL COMMENTS** | Accept all the changes. |