PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

This paper was submitted to the JECH but declined for publication following peer review. The authors addressed the reviewers’ comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

ARTICLE DETAILS

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>Real-life evaluation of European and American high-risk strategies for primary prevention of cardiovascular disease in patients with first myocardial infarction</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Falk, Erling; Mortensen, Martin</td>
</tr>
</tbody>
</table>

VERSION 1 - REVIEW

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Ben Hudson</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Otago</td>
<td>New Zealand</td>
</tr>
</tbody>
</table>

| REVIEW RETURNED   | 11-Jul-2014 |

| GENERAL COMMENTS | This paper provides an interesting retrospective view on the performance of several commonly used cardiovascular disease prevention guidelines amongst a group of patients who have had their first MI. The authors found that, despite correlating well with each other in terms of predicting risk, the CVD guidelines differed in their placing of individuals above or below treatment thresholds. The authors suggest that their approach allows an estimation of the treatment guidelines’ treatment threshold sensitivity, however several aspects of their method may have lead to inaccuracy in this estimate.  

The authors excluded patients with diabetes from their analysis. This excluded 92 potential participants who might have been expected to have higher estimated cardiovascular risk. Their inclusion in the analysis could have improved the apparent discriminating effect of the CVD risk tools.  

Statin users were similarly excluded. This was a smaller group (40), but these patients might also have represented individuals who, before their MIs, had been found to have high estimated CV risk (and had therefore been prescribed a statin). Their exclusion could similarly have lead to underestimation of the ability of the CVD guidelines to identify those at highest CVD risk.  

The CVD guidelines examined in the study include outcomes in addition to MI (e.g. stroke, fatal cardiovascular disease, TIA, angina). Ignoring these outcomes introduces further inaccuracy to estimates of the guidelines’ abilities to discriminate between high and low risk individuals. |
The investigators had to rely on collecting some risk factor measurements after participants' MIs. Whilst this is likely to have little effect on cholesterol levels, I suspect the same may not be true for blood pressure. Some participants' BPs were recorded just before their discharge or at their first rehabilitation clinic visit. It is highly likely that these participants would have begun one or more BP-lowering medications during their admission, for example ACE-inhibitors and beta-blockers, and therefore the BPs used in calculating the pre-MI CVD risk estimates are likely to be lower than was in fact the case, again leading to an underestimate of participants' pre-MI CVD risk. The authors provide some reassurance on this point by suggesting that there is little difference in the pre and post-MI BPs of “99 patients in whom paired values were available for comparison” and I assume these were patients included in this study, but this is not clear and it would be useful to provide some figures to illustrate the degree of similarity or difference.

The ideal method for estimating the ability of CVD risk guidelines to identify those at highest risk of CVD is a prospective cohort study, and when measured in this way their performance may be even worse than was found in the current study. In their commentary on the 2009 validation of QRISK2 (1) for example, Jackson et al pointed out that whilst 10% of men were classified as high risk, only 30% of subsequent CV events occurred in the high risk group, and the capture of CV events was even worse amongst women. (2)

Inconsistency in different CVD guidelines' placing of individuals above or below treatment thresholds has been demonstrated previously and it would be useful to compare the current study's findings with such work. For example, an analysis of 25 different CVD risk calculators' performance with 128 hypothetical patients found that 41% of the “patients” were assigned across categories of risk from low to high. (3)

The abstract is a little misleading in referring to 393 patients hospitalised with a first MI but not giving the actual number of patients included in the analysis (247). This should be changed.

The authors point out correctly that they cannot determine the CVD guidelines' specificity but that this can be deduced from a representative ROC curve. Given the limitations I have outlined above, any estimation of specificity is likely to be very imprecise and I think this claim should be further qualified or removed.


REVIEWER
Baris Gencer
Cardiology Division
Geneva University Hospitals
GENERAL COMMENTS
This a very interesting and well-written manuscript on the application of different cardiovascular prevention strategies in a population of patients with myocardial infarction in Denmark. The authors assessed the (1) calibration of different predictive models assessing the accuracy of high-risk estimation and (2) the proportion of patients who would require statin therapy based on recommended decision thresholds. The new 2013 ACC/AHA guidelines were able to classify all patients with MI in the high-risk category with the indication of statin therapy in primary prevention. In comparison, the low-risk SCORE equation was able to detect only a minority of patients with MI who would potentially benefit of statin therapy before the MI event. All the strategies were especially unable to detect the risk in women younger than 60 years old. However, the Reviewer has some major concerns as described below. Especially the limited sample size, and the focus on only untreated MI might be a potential source of misinterpretation of the most cost-effective prevention strategy in an European population.

Major Comments

1-The title « …In Europe and United States » might be a subject of controversy as the population of the study is from Danemark. In addition, the terminology « reality check » could be clarified in the article. A possible alternate title « Real-life evaluation of European and American high-risk strategies for the primary prevention of cardiovascular disease in patients with myocardial infarction in Denmark. »

2-The potential selection bias should be mentionned and discussed. The studied population is not representative of all patients with myocardial infarction, especially the sample size is very limited and the recruiting centers poorly described. The selection of patients should be described : (1) screening period, (2) name of hospitals, (3) description of the source documents (medical records, administrative dataset, CRF collection) and (4) the definition of myocardial infarction (e.g. criteria for ECG, cardiac enzymes, angiography, STEMI versus all acute coronary syndromes)

3-The calibration of model was assessed considering only the « cases » and not « non-cases ». The Reviewer has a major concern on a potential biased conclusion of such analysis as the model with the highest sensitivity will be « the winner » in contrast to the model with the highest specificity. In addition, the cross-sectional design in another limitation in the calibration of model. For instance, nearly 100% of men older than 60 years old are eligible to statin therapy according top new 2013 AHA/ACC guidelines. This expected impact of applying such strategy should be discussed in term of cost-effectiveness in the manuscript (see comment 5).

4-The patients treated with statin were excluded from main analysis. Thus limits the potential discussion in the manuscript on the adequacy of statin treatment. In the paper, the analysis was done only in untreated cases, discussion should also consider that a substantial proportion of patients was treated with statin prior to MI.

5-Analyzing description of applying prevention strategies. The authors should describe the issue of overtreatment using new
AHA/ACC guidelines. No study has compared in an European population the cost-effectiveness of different strategies in a randomized controlled trial. Thus the answer of the most optimal risk-based prevention strategy is still unresolved.

6-The authors should also add some points in the discussion regarding the « paradox » of prevention » described by Geoffrey Rose. The patients with the lowest risk will represent the most number of « cases » compared to those with the highest risk. This implies that “population strategies” which focus on reducing the risk of those already at low or moderate risk will often be more effective than strategies focusing on “high risk” individuals at improving population health. For example prevention strategies on lifestyles would have a greater impact in the whole population.

Minor comments:

Spell the abbreviations in the footnotes of Table 2 and 3.

Reviewer: Michael Blaha
Johns Hopkins
Baltimore, USA

I have conducted analyses of the discrimination and calibration of competing risk scores in the past.

Review returned: 23-Jul-2014

General comments:

3. The study design is appropriate to answer the study question (sensitivity of the risk scores/associated cutpoints for “detecting” a future MI). However the dataset is unable to comment on discrimination and calibration, which the authors do make reference to throughout the paper (as secondary points).

6. Myocardial infarction should be better defined - any MI, STEMI vs. NSTEMI, exclusion of peri-procedural MI, etc?

7. The statistics for the core study question are appropriate. However, the statistics for the comments on discrimination and calibration are suspect.

12. I do think the limitations should be expanded. The authors are looking at just one side of the question of appropriate use of statins - sensitivity. By this logic, the treat all approach would by definition be best - perfect sensitivity. The limitations of examining just sensitivity should be discussed in more detail.

Version 1 – Author Response

Reviewer: Ben Hudson

This paper provides an interesting retrospective view on the performance of several commonly used cardiovascular disease prevention guidelines amongst a group of patients who have had their first MI. The authors found that, despite correlating well with each other in terms of predicting risk, the CVD guidelines differed in their placing of individuals above or below treatment thresholds. The authors
suggest that their approach allows an estimation of the treatment guidelines’ treatment threshold sensitivity, however several aspects of their method may have lead to inaccuracy in this estimate.

The authors excluded patients with diabetes from their analysis. This excluded 92 potential participants who might have been expected to have higher estimated cardiovascular risk. Their inclusion in the analysis could have improved the apparent discriminating effect of the CVD risk tools.

Response: We appreciate all the comments from Dr. Ben Hudson. We wanted to estimate the detection rate (sensitivity) of the risk equations and decision thresholds recommended by the guidelines and, consequently, excluded patients who would not be eligible for risk assessment before their first MI. Patients with prevalent CVD or already using statins would not be eligible for risk assessment, and diabetic patients are not included in the target population for risk assessment defined by the ESC and ATP III guidelines. To clarify this issue we have added the following sentence under Study population: “(diabetic patients do not belong to the target population for risk assessment defined by the ESC and ATP III guidelines)”.

Statin users were similarly excluded. This was a smaller group (40), but these patients might also have represented individuals who, before their MIs, had been found to have high estimated CV risk (and had therefore been prescribed a statin). Their exclusion could similarly have lead to underestimation of the ability of the CVD guidelines to identify those at highest CVD risk.

Response: Theoretically, the reviewer is right. However, the potential impact on our results of excluding the relatively few statin users is limited. Among the 40 patients receiving statin for primary prevention only 26 were between 40 and 75 years of age (study population). Furthermore, although the 26 statin users aged 40-75 were slightly older than non-statin users aged 40-75 (65y vs 62y, p=0.06), they were more often non-smokers (33% vs 53%, p=0.05), had lower total cholesterol (4.3 mmol/l vs 5.3 mmol/l, p<0.0001) and comparable systolic blood pressures (135 mmHg vs 137 mmHg, p=0.72). Thus, statin-users did not have a higher SCORE risk compared with non-statin users based on their current risk factors (4.9% vs 5.6%, p=0.86). However, we believe that this might reflect that statin-users were more responsive to other risk lowering initiatives such as smoking cessation. For this reason, inclusion of statin users in our study would not improve the detection rate of the guidelines. To clarify this issue, we have added the low number of statin users (n=26) in the study population (age 40-75) in the Limitation section.

The CVD guidelines examined in the study include outcomes in addition to MI (e.g. stroke, fatal cardiovascular disease, TIA, angina). Ignoring these outcomes introduces further inaccuracy to estimates of the guidelines’ abilities to discriminate between high and low risk individuals.

Response: To avoid confusion we have now replaced “CVD” by “MI” throughout the paper, including the abstract, when referring to our own results.

The investigators had to rely on collecting some risk factor measurements after participants’ MIs. Whilst this is likely to have little effect on cholesterol levels, I suspect the same may not be true for blood pressure. Some participants’ BPs were recorded just before their discharge or at their first rehabilitation clinic visit. It is highly likely that these participants would have begun one or more BP-lowering medications during their admission, for example ACE-inhibitors and beta-blockers, and therefore the BPs used in calculating the pre-MI CVD risk estimates are likely to be lower than was in fact the case, again leading to an underestimate of participants’ pre-MI CVD risk. The authors provide some reassurance on this point by suggesting that there is little difference in the pre and post-MI BPs of “99 patients in whom paired values were available for comparison” and I assume these were patients included in this study, but this is not clear and it would be useful to provide some figures to illustrate the degree of similarity or difference.
The mean systolic blood pressure (±SD) measured before MI was 139.4 (±20.3) mmHg (n=103), after MI 137.1 (±19.0) mmHg (n=293, p=0.28).*

The ideal method for estimating the ability of CVD risk guidelines to identify those at highest risk of CVD is a prospective cohort study, and when measured in this way their performance may be even worse than was found in the current study. In their commentary on the 2009 validation of QRISK2 (1) for example, Jackson et al pointed out that whilst 10% of men were classified as high risk, only 30% of subsequent CV events occurred in the high risk group, and the capture of CV events was even worse amongst women. (2)

Response: We agree. The two papers from BMJ 2009 are very interesting but should be interpreted cautiously now NICE has lowered the high risk threshold from 20% to 10% which of course will increase the sensitivity of this threshold substantially.

Inconsistency in different CVD guidelines' placing of individuals above or below treatment thresholds has been demonstrated previously and it would be useful to compare the current study's findings with such work. For example, an analysis of 25 different CVD risk calculators' performance with 128 hypothetical patients found that 41% of the "patients" were assigned across categories of risk from low to high. (3)

Response: A major strength of our study is that we are describing results obtained in a contemporary cohort of real-life patients seen in clinical practice today, using risk prediction tools and decision thresholds recommended in current guidelines. Jackson et al wrote a thoughtful editorial in which they discussed potential limitations of the study by Allan et al, including problems related to (re)calibration. We now refer to this editorial (ref 27) in the Discussion under the subheading "Limitations of the traditional high-risk strategy based on prospective cohort studies".

The abstract is a little misleading in referring to 393 patients hospitalised with a first MI but not giving the actual number of patients included in the analysis (247). This should be changed.

Response: It has been changed.

The authors point out correctly that they cannot determine the CVD guidelines' specificity but that this can be deduced from a representative ROC curve. Given the limitations I have outlined above, any estimation of specificity is likely to be very imprecise and I think this claim should be further qualified or removed.

Response: It has been removed.

Reviewer: Baris Gencer

This a very interesting and well-written manuscript on the application of different cardiovascular prevention strategies in a population of patients with myocardial infarction in Denmark. The authors assessed the (1) calibration of different predictive models assessing the accuracy of high-risk estimation and (2) the proportion of patients who would require statin therapy based on recommended decision thresholds. The new 2013 ACC/AHA guidelines were able to classify all patients with MI in the high-risk category with the indication of statin therapy in primary prevention. In comparison, the low-risk SCORE equation was able to detect only a minority of patients with MI who would potentially benefit of statin therapy before the MI event. All the strategies were especially unable to detect the risk in women younger than 60 years old. However, the Reviewer has some major concerns as described below. Especially the limited sample size, and the focus on only untreated MI might be a
potential source of misinterpretation of the most cost-effective prevention strategy in an European population.

Major Comments:

1-The title « …In Europe and United States » might be a subject of controversy as the population of the study is from Danemark. In addition, the terminology « reality check » could be clarified in the article. A possible alternate title « Real-life evaluation of European and American high-risk strategies for the primary prevention of cardiovascular disease in patients with myocardial infarction in Denmark. »

Response: Many thanks for the comments and the proposed title. The title has now been changed to “Real-life evaluation of European and American high-risk strategies for primary prevention of cardiovascular disease in patients with first myocardial infarction”.

2-The potential selection bias should be mentioned and discussed. The studied population is not representative of all patients with myocardial infarction, especially the sample size is very limited and the recruiting centers poorly described. The selection of patients should be described: (1) screening period, (2) name of hospitals, (3) description of the source documents (medical records, administrative dataset, CRF collection) and (4) the definition of myocardial infarction (e.g. criteria for ECG, cardiac enzymes, angiography, STEMI versus all acute coronary syndromes).

Response: With the study design used, including all consecutive, contemporary patients with a first MI seen clinically, potential selection bias regarding the performance of clinical guidelines should not be a concern. The sample size is, in fact, not “very limited”. To address this question, we have added the following sentence and a supporting reference in the Discussion section: “To put it into perspective, our reality check included 247 patients hospitalized with a first MI, in the JUPITER trial only 62 nonfatal MI were observed among 8901 placebo patients during nearly 2 years of follow-up (31).” The requested information about the study population has been added under the subheading “Study population”. Briefly, the patients included in this study were patients admitted to the departments of cardiology/medicine at Aarhus University Hospital and the Regional Hospitals in Herning and Randers with a first MI between January 1 through December 31 in 2011. The universal definition of MI is implemented in Denmark (Thygesen K, Alpert JS, White HD; Joint ESC/ACCF/AHA/WHF Task Force for the Redefinition of Myocardial Infarction. Universal definition of myocardial infarction. Eur Heart J. 2007;28:2525-38), and the patients were identified via hospital registers using ICD-10 codes I21.0 through I21.9. Thus, our study includes both STEMI and non-STEMI patients.

3-The calibration of model was assessed considering only the « cases » and not « non-cases ». The Reviewer has a major concern on a potential biased conclusion of such analysis as the model with the highest sensitivity will be « the winner » in contrast to the model with the highest specificity. In addition, the cross-sectional design in another limitation in the calibration of model. For instance, nearly 100% of men older than 60 years old are eligible to statin therapy according to new 2013 AHA/ACC guidelines. This expected impact of applying such strategy should be discussed in term of cost-effectiveness in the manuscript (see comment 5).

Response: We agree. There always is a trade-off between sensitivity and specificity of diagnostic (and prognostic) tests. However, our study design only allows assessment of sensitivity, not specificity and cost-effectiveness. To clarify these issues we have added the following text in the Discussion under the subheading “Limitations of the present study”: “In patients with a first MI, only the detection rate (sensitivity) of decision thresholds can be determined, not the specificity and risk of overtreatment. However, if a decision threshold captures only a minority of those it was intended to identify, its utility may be questioned. Given that the 2013 ACC/AHA and the 2014 NICE/UK
guidelines lowered the threshold for primary prevention with statin based on careful risk-benefit and cost-effectiveness considerations.\[3,4,33\] the appropriateness of the much lower sensitivity of the SCORE-based treatment threshold recommended for use in many high-income European countries\[1\] deserves to be reconsidered."

4-The patients treated with statin were excluded from main analysis. Thus limits the potential discussion in the manuscript on the adequacy of statin treatment. In the paper, the analysis was done only in untreated cases, discussion should also consider that a substantial proportion of patients was treated with statin prior to MI.

Response: As mentioned in our response to Dr. Ben Hudson (please see above), the potential impact on our results of excluding the relatively few statin users is limited (only 26 patients in the study population used statin).

5-Analyzing description of applying prevention strategies. The authors should describe the issue of overtreatment using new AHA/ACC guidelines. No study has compared in an European population the cost-effectiveness of different strategies in a randomized controlled trial. Thus the answer of the most optimal risk-based prevention strategy is still unresolved.

Response: It is indeed important questions, but it was not the purpose of our study to answer these questions. However, we have added the following two sentences in the Discussion under the subheading "US guideline": "Concerns have been raised about the potential risk of overtreatment.\[23\]" and "A recent review concluded that the new recommendations for statin therapy “generally meet societal acceptable levels of cost-effectiveness."\[24\]."

6-The authors should also add some points in the discussion regarding the « paradox » of prevention described by Geoffrey Rose. The patients with the lowest risk will represent the most number of « cases » compared to those with the highest risk. This implies that “population strategies” which focus on reducing the risk of those already at low or moderate risk will often be more effective than strategies focusing on “high risk” individuals at improving population health. For example prevention strategies on lifestyles would have a greater impact in the whole population.

Response: Cheap, effective and safe risk reducing drugs were not available at the time Geoffrey Rose 30 years ago described the “prevention paradox”. Regarding these important questions, we don’t think our study can provide guidance beyond the view expressed in the amended Conclusion on page 10: "In the US and UK, a treatment threshold based on risk-benefit and cost-effectiveness considerations \[3,4,24,33\] has now been defined, leading to a wider eligibility for primary prevention with statin therapy".

Minor comments:

Spell the abbreviations in the footnotes of Table 2 and 3.

Response: Thanks, abbreviations have now been spelled out.

Reviewer: Michael Blaha

The study design is appropriate to answer the study question (sensitivity of the risk scores/associated cutpoints for “detecting” a future MI). However the dataset is unable to comment on discrimination and calibration, which the authors do make reference to throughout the paper (as secondary points).

Myocardial infarction should be better defined - any MI, STEMI vs. NSTEMI, exclusion of peri-procedural MI, etc?
Response: All the comments from Dr. Michael Blaha are highly appreciated. Nearly all our comments on discrimination and calibration have been deleted (further explained below). As clarified in our answer to Dr. Baris Gencer (please see above), the study population and MI are now better defined. We used the universal definition of MI (Thygesen K, Alpert JS, White HD; Joint ESC/ACCF/AHA/WHF Task Force for the Redefinition of Myocardial Infarction. Universal definition of myocardial infarction. Eur Heart J. 2007;28:2525-38), which includes both STEMI and non-STEMI. Peri-procedural MI is unlikely because patients with known ASCVD were excluded from the study population.

The statistics for the core study question are appropriate. However, the statistics for the comments on discrimination and calibration are suspect.

Response: All comments on discrimination and calibration based on our dataset have been deleted.

I do think the limitations should be expanded. The authors are looking at just one side of the question of appropriate use of statins - sensitivity. By this logic, the treat all approach would by definition be best - perfect sensitivity. The limitations of examining just sensitivity should be discussed in more detail.

Response: We have expanded the Limitation section, and the last few sentences now read as follows: “In patients with a first MI, only the detection rate (sensitivity) of decision thresholds can be determined, not the specificity and risk of overtreatment. However, if a decision threshold captures only a minority of those it was intended to identify, its utility may be questioned. Given that the 2013 ACC/AHA and the 2014 NICE/UK guidelines lowered the threshold for primary prevention with statin based on careful risk-benefit and cost-effectiveness considerations,[3,4,33] the appropriateness of the much lower sensitivity of the SCORE-based treatment threshold recommended for use in many high-income European countries[1] deserves to be reconsidered.”

Furthermore, we have added the following two sentences in the Discussion under the subheading "US guideline": "Concerns have been raised about the potential risk of overtreatment.[23]" and “A recent review concluded that the new recommendations for statin therapy “generally meet societal acceptable levels of cost-effectiveness.”[24]."

This is a very interesting and important paper. The question of the best risk score for allocating medications in primary prevention is of paramount importance. The new ACC/AHA guidelines bring about a host of new questions, and the authors have squared in on the comparison of the “detection rate” of each of the guidelines.

It should be noted that this looking at just one side of the risk prediction coin. Detection rate, or sensitivity, will improve any time the threshold to treat is lowered. Thus is has little to do with the risk score, only the threshold chosen. As such, the treat all approach (which is advocated by some) will always be superior from this perspective - it will by definition have perfect sensitivity.

Of course, this is a very simplistic approach to risk prediction and medication allocation. Most prefer a treat to risk (not treal all) perspective because this will best balance the benefits with potential risks. Thus treat all will expose many to medications without the prospect of benefit. I.e. - when sensitivity goes up, specificity goes down.

As this paper is only looking at cases (everyone has the outcome), it is impossible to comment on discrimination and calibration. The authors comment on discrimination (the rank order), but of course this is highly correlated across risk scores because they all use about the same variables. I would recommend removing any comment on discrimination/rank-order from the paper. This should be reserved for large population based studies where not everyone experiences the outcome.
Response: We appreciate all these thoughtful comments. We understand that it is a simplistic approach but still believe it may provide important information, complementary to that obtained from prospective cohort studies as discussed under the subheading “Limitations of the traditional high-risk strategy based on prospective cohort studies”. We agree that our study provides no data on discrimination and specificity and have deleted the paragraph discussing these aspects.

Likewise, I would recommend removing any comments on calibration from the discussion. There is no way to comment on true calibration from this study. I did not follow some of the back of the envelope calculations presented in the discussion section of this paper, particularly as it pertains to comments on calibration of the new ACC/AHA risk score. Once again, this can only be answered in a prospective cohort study.

Response: We appreciate this comment and have followed the advice. The text under the subheading “US guideline” in the Discussion was removed and replaced by the following much shorter version:

“With the release of the 2013 ACC/AHA guideline, a new risk calculator based on PCE was introduced.[3] We were not able to access calibration of PCE in our study population, but recent data indicates that PCE is reasonable well calibrated (similar to SCORE) in a UK “low-risk” population.[20] This observation provides a reasonable background for comparing PCE directly with SCORE and QRISK in a European country classified as “low-risk” (Figure 3 and 5).

We estimated and compared predicted risk and found that a PCE risk of 7.5% corresponded to an ATPIII risk of ~10% in men and ~4% in women (Figure 1, Table 3), documenting that the bar for primary prevention with statin therapy deliberately was lowered by the 2013 ACC/AHA guideline, especially in women.[21,22] Concerns have been raised about the potential risk of overtreatment.[23] In patients with first MI who were 60 to 75 years of age, the sensitivity of the new class I recommendation for statin therapy (PCE ≥7.5%) was 100% in men and 85% in women (Figure 2). The new cut-point for treatment was established based on risk-benefit considerations alone.[2,3] cost-effectiveness of fixed-dose not target driven statin therapy was not questioned.[22] A recent review concluded that the new recommendations for statin therapy “generally meet societal acceptable levels of cost-effectiveness.”[24] The ACC/AHA guidelines are expected to increase the number of people eligible for primary prevention with statins in the US substantially.[25]”

Minor comments:

1. I would refrain from calling 10-year risk "short term". 10-year risk remains the parlance of clinical risk prediction. I would just call it 10-year risk, without the moniker "short term"

Response: “Short term” has been deleted.

2. In the introduction, "accuracy" is equated with "calibration". This is not correct. Accuracy is reflected in 2 distinct concepts - discrimination and calibration.

Response: In the introduction, “accuracy” has been deleted.

3. I agree that the lack of perfect measurements of BP and cholesterol probably don’t impact the results, given that risk scores are mostly age and gender. However, this raises the question - couldn’t this study mostly be accomplished with a series of made up patients? This is really a modeling exercise looking at sensitivity across varying thresholds, and essentially any patient population could be used to make this point.
Response: We wanted to estimate the sensitivity of the guideline-defined treatment threshold in contemporary, real-life patients representative of those seen in clinical practice today (actual distribution of age, sex, smoking etc). It is difficult to model in hypothetical patients.

I think this is a very worthy exercise and I would highly appreciat the opportunity to comment on revisions of this manuscript. This topic will be of high interest to the journal's readers.

**VERSION 2 – REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Baris Gencer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Geneva University Hospitals</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>18-Aug-2014</td>
</tr>
</tbody>
</table>

- The reviewer completed the checklist but made no further comments

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Michael Blaha</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Johns Hopkins, USA</td>
</tr>
<tr>
<td></td>
<td>No conflicts</td>
</tr>
<tr>
<td>REVIEW RETURNED</td>
<td>12-Sep-2014</td>
</tr>
</tbody>
</table>

- The reviewer completed the checklist but made no further comments