PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Can traditional risk factors explain the higher risk of cardiovascular disease in South Asians compared to Europeans in Norway and New Zealand? Two cohort studies
AUTHORS	Rabanal, Kjersti; Meyer, Haakon; Tell, Grethe; Igland, Jannicke; Pylypchuk, Romana; Mehta, Suneela; Kumar, Bernadette; Jenum, Anne; Selmer, Randi; Jackson, Rod

VERSION 1 - REVIEW

REVIEWER	David Webb University of Leicester, UK
REVIEW RETURNED	31-Mar-2017

GENERAL COMMENTS	1) The abstract conclusion, that preventing dyslipidaemia in south Asians may help reduce their excess risk is probably true but not substantiated by the findings of this paper or its main goal of establishing whether risk is predominantly accounted for with known factors. 2) Reporting two entirely independent datasets in one paper is unusual and in my opinion an unnecessary distraction. I am not convinced by the “external validity” argument as the authors state themselves that the cohorts are very different and also cite other published work with similar findings. Therefore the authors should consider removing the Norwegian data or providing a better explanation for its inclusion. 3) A “prioritising algorithm” was used to code for ethnicity status. The details of this need to be more comprehensively described in the supplementary section if necessary. 4) Missing data is an inevitable issue for clinical datasets such as PREDICT. Please expand the limited text covering this important limitation and explain in greater detail how it was handled. 5) The statement “all analyses were stratified by sex and ethnicity” in the statistical analysis section of the methods following on directly after a sentence about ethnicity as the exposure variable is confusing p10 line28. Please amend accordingly 6) By virtue of the source screening criteria, South Asians in the New Zealand cohort are some 7 years younger than their White European counterparts. Whilst attempts have been made to statistically adjust for this difference, this remains an issue which needs to be acknowledged as increasing age is the strongest determinant of cardiovascular disease. 7) Inconsistency in tabulated results reporting – better to stratify table one by cohort then gender. 8) There are a number of ways of determining the relative impact of individual risk factors and therefore residual unaccounted for risk.

	The authors may which to look at mediation analysis. 9) Please explain how the number of events data in Table 3 relate to the description of events in the text p14 line12-17. What do the asterisks refer to in table 3? 10) The authors are a little vague about mortality data eg. “Non-cardiovascular causes of death were excluded”. Exact numbers and causes of death where known would be interesting and readily obtainable via the linkage source statistics described in the methodology. 11) Similarly it would be interesting to breakdown the composite CVD term eg. Is the excess in South Asians primarily coronary heart disease related in line with existing literature? 12) Whilst the authors acknowledge that populations derived from health screening initiatives are generally not representative of background, there is some evidence that the uptake of such programmes in migrant or ethnic minority groups is even lower. This may be troublesome when trying to directly compare risks that are picked up by such programmes. 13) An important limitation of this work is the lack of available medication data. Discrepancies in prescription of or compliance with cardiovascular protection medication such as statins and antihypertensives could account for some of the residual risk and should be discussed.

REVIEWER	Jack Tu ICES Canada
REVIEW RETURNED	23-Apr-2017

GENERAL COMMENTS	1. Although the authors’ focus is not to compare findings between the cohorts, the cohorts are sufficiently different (e.g., in defining South Asians, data sources, population represented, exclusion criteria, length of follow-up, outcome definitions) that by combining into 1 paper, comparisons are inherent yet less meaningful than what could be learned from a more in depth analyses of the cohorts alone. 2. A major issue with the analyses is the rationale for the order of risk factor adjustment in the multivariable analysis (Table 3) should be provided since changing the order in which the risk factors were added to the models could change the results and interpretation. As is, the most useful results are the age- and fully-adjusted models. Additionally, the age-adjusted hazard ratios in Table 2 show smoking generally had the greatest effect on CV risk among the risk factors studied, and thus, could be considered first. 3. It would be helpful to report unadjusted hazard ratios and age-adjusted/standardized CV event rates risks by sex and ethnicity in the main paper for comparisons. 4. Sensitivity analyses – excluding those on medications likely excludes those at greatest risk of CVD. Since the characteristics of those excluded by ethnicity/cohort are not reported, there is potential for differential bias. Did the authors consider adjusting for medication use as an alternative? 5. While a breakdown of the proportion of CV events by specific conditions was not provided and the small number of events among the South Asian populations may have necessitated a broad definition of the CV outcome, a sensitivity analysis restricting to the

	major CV events (i.e. IHD, stroke) should be considered as the conditions included have varying risk factors. Minor points: 6. The flow of the paper could be improved if there was consistency in the order of describing the methods and presenting the results with respect to the 2 cohorts, i.e. New Zealand followed by Norway or vice versa 7. Although the ICD codes for outcomes are provided, it would be more helpful if the conditions associated with these codes were included. 8. Although the outcomes were assessed after risk factor measurement, the study design may be better described as a retrospective +/- cohort study rather than prospective study as the study appears to have been planned/conducted post-data collection.

VERSION 1 – AUTHOR RESPONSE

REVIEWER 1

Reviewer Name: David Webb
Institution and Country: University of Leicester, UK

COMMENTS FROM REVIEWER 1
The authors report Cox proportional hazards for cardiovascular events (and mortality) in the South Asian diaspora of two predominantly White European datasets. They conclude that after 7-8 and 2-3 years of follow up and 743 and 2654 events respectively, the risk of cardiovascular disease is significantly higher in South Asians compared with “indigenous” White Europeans and can only partly be explained by so-called traditional risk factors.
Whilst largely confirmatory, these findings are important and consolidate our understanding of common disease patterns in too often overlooked BME groups.
The paper is clear in its intentions, its methodology mostly appropriate and its results convincing. I felt the discussion lacked depth in exploring possible explanations for the findings but is otherwise erudite and balanced.
I invite the reviewers to respond to the following suggestions and recommendations which will hopefully improve the paper.

1) The abstract conclusion, that preventing dyslipidaemia in south Asians may help reduce their excess risk is probably true but not substantiated by the findings of this paper or its main goal of establishing whether risk is predominantly accounted for with known factors.

REPLY: We agree that this is not a direct conclusion of our findings. However, it is an implication of our conclusion and since BMJ Open recommends that the abstract conclusion should contain “primary conclusions and their implications”, we have decided to keep it. However, we added the word “therefore” to better show the linkage between our conclusion (the first sentence) and its implication (the second sentence). See page 3.

2) Reporting two entirely independent datasets in one paper is unusual and in my opinion an unnecessary distraction. I am not convinced by the “external validity” argument as the authors state themselves that the cohorts are very different and also cite other published work with similar findings.

Therefore the authors should consider removing the Norwegian data or providing a better explanation for its inclusion.

REPLY: We understand the reviewers point, but having two different datasets in one paper can also be regarded as an advantage. We found the same patterns in the two cohorts regarding the contribution of risk factors for the excess risk of CVD in South Asians. We therefore think that the "external validity" point is relevant and especially since the two cohorts are so different.
The main reason for including both cohorts, however, is the lack of studies assessing the prospective association between risk factors and CVD among South Asians. Previous studies have mainly been cross-sectional. This makes us hesitant to remove the Norwegian data because we think it adds valuable information to the research field.
Please also see our response to reviewer 2 (the general comment which comes before the numbered comments).

3) A "prioritising algorithm" was used to code for ethnicity status. The details of this need to be more comprehensively described in the supplementary section if necessary.

REPLY: We have now added an additional supplementary file (Supplementary_Revised VIEW Ethnicity Protocol 17December2015.docx) that describes this in more detail.

4) Missing data is an inevitable issue for clinical datasets such as PREDICT. Please expand the limited text covering this important limitation and explain in greater detail how it was handled.

REPLY: As stated in the methods section, only complete cases were included in the analyses. For the PREDICT data, there were almost no missing (0,01%) on the four major risk factors that we studied because it was compulsory for the general practitioner to fill in these four risk factors in the PREDICT template as they were part of the prediction algorithm. We have now added a sentence to clarify this on page 17.
There was a bit more missing in the Norwegian cohort, however, and we have added a few sentences about this limitation in the discussion section (page 18).

5) The statement "all analyses were stratified by sex and ethnicity" in the statistical analysis section of the methods following on directly after a sentence about ethnicity as the exposure variable is confusing p10 line28. Please amend accordingly

REPLY: We agree with the reviewer and have amended the statement (page 10).

6) By virtue of the source screening criteria, South Asians in the New Zealand cohort are some 7 years younger than their White European counterparts. Whilst attempts have been made to statistically adjust for this difference, this remains an issue which needs to be acknowledged as increasing age is the strongest determinant of cardiovascular disease.

REPLY: We agree that age is the most important determinant and that age differences between the ethnic groups generally make comparing complicated. However, since we have adjusted for age in the Cox regression we do mean that the ethnic groups are comparable (within the cohorts), and that our results are valid despite the age differences. We have added a sentence about this issue in the discussion section of the manuscript, page 17.

7) Inconsistency in tabulated results reporting – better to stratify table one by cohort then gender.

REPLY: We are a bit confused about what the reviewer means since Table 1 is already stratified by cohort. Regarding the other tables, we initially tried different solutions before eventually stratifying by gender with an "under-stratification" by cohort because (in our opinion) it made comparisons easier

and the tables more readable. We are, however, happy to do otherwise if advised more clearly.

8) There are a number of ways of determining the relative impact of individual risk factors and therefore residual unaccounted for risk. The authors may which to look at mediation analysis.

REPLY: We have not performed a formal mediation analyses because we are in a situation with several risk factors/mediators that are related to each other. This makes it difficult to separate the different indirect effects [1]. We have, however, added two sentences about how much the estimates for the excess risk in South Asians are reduced when adjusting for the four risk factors (see page 15-16).

9) Please explain how the number of events data in Table 3 relate to the description of events in the text p14 line 12-17. What do the asterisks refer to in table 3?

REPLY: We thank the reviewer for noticing the inconsistency between Table 3 and the text on page 14 (now page 13) since there were a typing error in Table 3 (the 8631 events among women were in fact only 863.). We have now corrected this.
The numbers of events still differ somewhat between the description in the text on page 13 and in Table 3. This is because the text on page 13 describes the numbers of events for the whole datasets, while persons with missing on any of the four risk factors are excluded from Table 3. For the Norwegian cohort, there were 700 events among persons with complete information on the four risk factors (which means that 43 events occurred among persons with missing information).
Since there were few missing in PREDICT, the numbers of events are the same in Table 3 and in the text on page 13.
Regarding the asterisk, we forgot to remove it after deleting a footnote, but have removed it now.

10) The authors are a little vague about mortality data eg. “Non-cardiovascular causes of death were excluded”. Exact numbers and causes of death where known would be interesting and readily obtainable via the linkage source statistics described in the methodology.

REPLY: First, we would like to comment that we did not exclude people if they died from non-cardiovascular causes, but they were censored in the Cox regression analyses. For CONOR, the number of people who were censored because they died from non-cardiovascular causes was 276 (17 South Asians and 259 Europeans). For PREDICT, the number was 961 (46 Indians and 915 Europeans). These total numbers are now added to the text (page 10).
For the datasets that we had access to, we only had information about the exact causes of death in CONOR, where cancer seemed to be the most common non-cardiovascular cause of death. Since we did not have this information in PREDICT, we have not included any information about non-cardiovascular causes of death in the manuscript. We are, however, happy to include it if advised to.

11) Similarly it would be interesting to breakdown the composite CVD term eg. Is the excess in South Asians primarily coronary heart disease related in line with existing literature?

REPLY: This is an interesting question and we have already looked at this, but decided to use the composite CVD as the outcome because of few endpoints among the South Asians and because we thought it would be too much to describe. As the reviewer suggests, and in line with existing literature, the excess risk in South Asians is especially high for coronary heart disease (CHD). We have now added some information about this in the manuscript, page 16.

See Table 1s in the supplementary file for editors only (called Supplementary files for Editors only.docx) showing CHD as the outcome (corresponding to Table 3).

12) Whilst the authors acknowledge that populations derived from health screening initiatives are generally not representative of background, there is some evidence that the uptake of such programmes in migrant or ethnic minority groups is even lower. This may be troublesome when trying to directly compare risks that are picked up by such programmes.

REPLY: This is a good point. For PREDICT, however, Maori , Pacific and Indian patients are over-represented in the cohort [2]. This is now commented in the manuscript, page 17. The over-representation of these high-risk ethnic groups is influenced by the New Zealand CVD guidelines recommendations for screening (as described in the methods section of the manuscript page 6).

13) An important limitation of this work is the lack of available medication data. Discrepancies in prescription of or compliance with cardiovascular protection medication such as statins and antihypertensives could account for some of the residual risk and should be discussed.

REPLY: We agree that this is an important limitation. For both cohorts, we had information about use of antihypertensives and lipid lowering drugs at baseline and we have now tried to adjust for baseline medication in the full-adjusted model in Table 3. These adjustments did not change the estimates much (see Table 5s in the supplementary file for editors only). A study from New Zealand comparing medication maintenance in secondary prevention also indicated that Indian patients were more likely to be maintained on
triple therapy compared to other ethnic groups[3]. Although the study did not look at primary prevention medication, it still indicates that Indians in New Zealand generally have good compliance to cardiovascular protection medication.
We have now added a few sentences about this issue on page 19 in the manuscript.

REVIEWER 2

Reviewer Name: Jack Tu
Institution and Country: ICES, Canada

COMMENTS FROM REVIEWER 2
This paper examines cardiovascular (CV) event rates in South Asians/Indians versus Europeans in Norway and New Zealand, and investigates the role of traditional cardiac risk factors in the differences observed. The Norwegian cohort is comprised of participants from 3 surveys conducted in Oslo, Norway with data linked hospitalization and death records for outcomes. The New Zealand PREDICT-CVD cohort comprises individuals who had a CV risk assessment by a general practitioner or nurse, and is also linked to national health datasets for outcomes. In multivariate Cox regression analyses, South Asians were found to be at higher risk of CVD than Europeans in both, but also had higher crude diabetes prevalence and mean TC/HDL ratios which contributed to their higher risk.
With increasing migration worldwide, the health of migrants is gaining interest. However, it is difficult to find much novelty in the study as several previous studies have reported greater risk of CVD among South Asian populations compared with their European counterparts in the same country (e.g. Fernando E, et al in Can J Cardio 2015), and the contribution of traditional rick factors to these differences. Some other comments follow.

Reply: First, we would like to thank the reviewer for his constructive comments. As the reviewer finds it difficult to see the novelty in our study, we would like to stress the scarcity of prospective data on the relationship between risk factors and later CVD among South Asians. This was underlined by

Meghna Ranganathan and Raj Bhopal in 2006 [4], and has recently been pointed out by Salim Yusuf and Philip Joseph [5], although the latter mentioned it explicitly for South Asians in the South Asian region. We have changed the wording and added a sentence regarding this in the introduction of the manuscript hoping to make it clearer to the reader (page 4).

1. Although the authors' focus is not to compare findings between the cohorts, the cohorts are sufficiently different (e.g., in defining South Asians, data sources, population represented, exclusion criteria, length of follow-up, outcome definitions) that by combining into 1 paper, comparisons are inherent yet less meaningful than what could be learned from a more in depth analyses of the cohorts alone.

REPLY: See our answer to reviewer 1, comment number 2.

2. A major issue with the analyses is the rationale for the order of risk factor adjustment in the multivariable analysis (Table 3) should be provided since changing the order in which the risk factors were added to the models could change the results and interpretation. As is, the most useful results are the age- and fully-adjusted models. Additionally, the age-adjusted hazard ratios in Table 2 show smoking generally had the greatest effect on CV risk among the risk factors studied, and thus, could be considered first.

REPLY: We have now added some information about our rationale for the order of risk factors adjustment in the methods section (page 10). In short, we first added the risk factors that were more prevalent among South Asians compared to Europeans (diabetes and TC/HDL ratio) because they could potentially explain the increased risk in South Asians. We then added the less prevalent risk factors systolic blood pressure and smoking.
Since the risk factors were on different scales, it cannot be concluded that smoking had the greatest effect (the risk factors are not comparable).
We have done additional analyses adding the risk factors in different orders to see whether the results and interpretation change, but the interpretation stays more or less the same (diabetes seem to have most to say for the excess risk in both cohorts and the lipids play a role for the excess risk in the Norwegian cohort). See Table 2s and Table 3s in the supplementary file for editors only where we show the results when adjusting for the risk factors in a different order than in the manuscript (Table 2s) and the results where all risk factors are in separate models only adjusted for age (Table 3s) - as alternatives to Table 3. We suggest to keep the table as it is.

3. It would be helpful to report unadjusted hazard ratios and age-adjusted/standardized CV event rates risks by sex and ethnicity in the main paper for comparisons.

REPLY: We are a bit confused about what the reviewer means here. We could have added unadjusted hazard ratios to Table 3, but we think it could make the table more complicated/difficult to read and that it could distract from the main message.
Hoping to answer some of the reviewers concerns, however, we have added the crude rates by sex and ethnicity (previously only shown in the appendices) in Table 2. Age adjusted event rates (in order to compare rates between the two cohorts) are not straight forward since the cohorts have different length of follow-up. We could have calculated standardized event rates, but this is a demanding task compared to the possible benefits in this context. We have previously presented age standardized event rates for different ethnic groups in the whole Norwegian population [6].

4. Sensitivity analyses – excluding those on medications likely excludes those at greatest risk of CVD. Since the characteristics of those excluded by ethnicity/cohort are not reported, there is potential for

differential bias. Did the authors consider adjusting for medication use as an alternative?

REPLY: We agree that we probably exclude the sickest when excluding those on medications. We have therefore done additional analyses adjusting for medication use as the reviewer suggests. Some information about this has been included in the manuscript, see page 10 and 14. In Table 4s in the supplementary file for editors only, we show the sensitivity analyses for Table 2. In Table 5s we show the full-adjusted model from Table 3 additionally adjusted for medication use. This did not change the estimates much either.

5. While a breakdown of the proportion of CV events by specific conditions was not provided and the small number of events among the South Asian populations may have necessitated a broad definition of the CV outcome, a sensitivity analysis restricting to the major CV events (i.e. IHD, stroke) should be considered as the conditions included have varying risk factors.

REPLY: We did consider this, but as the reviewer points out, we had limited power and therefore decided to use the broad CVD outcome. See Table 1s in the supplementary file for editors only, which is a table corresponding to Table 3 with CHD as the outcome. It was not meaningful to show separate results for stroke because of few events in some of the subgroups.

Minor points:

6. The flow of the paper could be improved if there was consistency in the order of describing the methods and presenting the results with respect to the 2 cohorts, i.e. New Zealand followed by Norway or vice versa

REPLY: We agree and have made changes so that the New Zealand cohort is described first (this also includes the discussion page 17-18). The text where we describe both cohorts simultaneously has not been changed. We have also changed the order of the two cohorts in Table 1 and Table 3 so that PREDICT is presented before CONOR.

7. Although the ICD codes for outcomes are provided, it would be more helpful if the conditions associated with these codes were included.

REPLY: We have now added a sentence about which conditions are associated with the listed ICD codes and thereby included in the CVD endpoint (see page 8).

8. Although the outcomes were assessed after risk factor measurement, the study design may be better described as a retrospective +/- cohort study rather than prospective study as the study appears to have been planned/conducted post-data collection.

REPLY: We thank the reviewer for making us reflect over this. However, we are hesitant to change the description to retrospective cohort since the risk factor data were collected before the disease events occurred and the data collection had a focus on cardiovascular risk factors (obvious so in the PREDICT dataset, but also in the Norwegian data [7]). Kenneth J. Rothman states that “a reasonable rule might be to describe a study as prospective if the exposure measurement could not be influenced by the disease, and retrospective otherwise” (page 96 in [8]). We therefore think that prospective cohort study is the best description for this study design.

References

1. VanderWeele, T.J., Mediation analysis: a practitioner's guide. Annual review of public health, 2016. 37: p. 17-32.

2. Wells, S., et al., Cohort Profile: The PREDICT Cardiovascular Disease Cohort in New Zealand Primary Care (PREDICT-CVD 19). International journal of epidemiology, 2015: p. dyv312.
3. Kerr, A., et al., Effect of age, gender, ethnicity, socioeconomic status and region on dispensing of CVD secondary prevention medication in New Zealand: The Atlas of Health Care Variation CVD cohort (VIEW-1). 2014.
4. Ranganathan, M. and R. Bhopal, Exclusion and inclusion of nonwhite ethnic minority groups in 72 North American and European cardiovascular cohort studies. PLoS medicine, 2006. 3(3): p. e44.
5. Yusuf, S. and P. Joseph, The epidemic of cardiovascular disease in South Asians: Time for action. American Heart Journal, 2017. 185: p. 150-153.
6. Rabanal, K.S., et al., Ethnic inequalities in acute myocardial infarction and stroke rates in Norway 1994-2009: a nationwide cohort study (CVDNOR). BMC Public Health, 2015. 15: p. 1073.
7. Kumar, B., et al., Ethnic differences in SCORE cardiovascular risk in Oslo, Norway. Eur J Cardiovasc Prev Rehabil, 2009. 16: p. 229 - 234.
8. Rothman, K.J., S. Greenland, and T.L. Lash, Modern epidemiology. 2008: Lippincott Williams & Wilkins.

VERSION 2 – REVIEW

REVIEWER	David Webb University of Leicester, UK
REVIEW RETURNED	25-Jun-2017

GENERAL COMMENTS	The authors have adequately addressed most of the reviewers concerns in a comprehensive and detailed rebuttal. I repeat this is important data and worthy of publication if only to reinforce some key messages in the few other reported prospective studies in migrant South Asians. In particular I was surprised to see that medication data was actually readily available for both cohorts and welcome its inclusion in revised analyses. I am also now reassured that missing data is not an issue although from my rough calculation about 2-3% of cases in the Norwegian cohort must have one or more risk factor missing. The authors have attempted to stratify the CVD events composite term although I am not sure why these results should be confined to the "editors only files". Strengths / Limitations are now more balanced and the discussion section in general is more convincing. All in all the manuscript is definitely improved but for the record I am still not convinced that consecutively reporting two very different cohorts in this way, even if they do say the same thing, enhances the readers experience or really adds much other than probable confusion.

REVIEWER	Jack Tu Institute for Clinical Evaluative Sciences, Sunnybrook Schumacher Heart Centre, University of Toronto Canada
REVIEW RETURNED	05-Jul-2017

GENERAL COMMENTS	The revised manuscript is improved over the original one. I think the authors should consider including as supplemental appendices, the analyses currently provided for the editors only.

VERSION 2 – AUTHOR RESPONSE

REPLY TO REVIEWER 1: We thank the reviewer for his comments and are happy to hear that he thinks the manuscript has been improved. We have now included the analyses where we adjust for medication at baseline in the appendices. We have also included the analyses where CHD is the outcome instead of broad CVD as part of the supplementary appendices.

REPLY TO REVIEWER 2: We thank the reviewer for his comments. We have now included the tables originally provided for the editors only as part of the supplementary appendices.