Author's response to reviews

Title: Risk factors for non-communicable disease and healthcare expenditure in employees with private health insurance presenting for health risk appraisal: A cross-sectional study

Authors:

Tracy L Kolbe-Alexander (tracy.kolbe-alexander@uct.ac.za)
Jaco Conradie (jconradie@humana.com)
Estelle V Lambert (vicki.lambert@uct.ac.za)

Version: 4 Date: 5 December 2013

Author's response to reviews: see over
Dr Santos
Editorial Office
Bio Med Central Editorial Office

Dear Dr Santos

I would like to thank the reviewer for taking the time to review the manuscript entitled, “Risk factors for non-communicable disease and healthcare expenditure in employees with private health insurance presenting for health risk appraisal: A cross-sectional study”.

Thank you for your comments, which as enhanced the quality of the manuscript. Please find a detailed response to the comments attached. The revisions are in blue text in the manuscript in order to assist the reviewer’s with locating the changes.

Furthermore, the abstract has been formatted and the declaration of non-financial competing interests has been included as suggested by the editor. The manuscript has been read by two fluent English (first language) speaking people.

We hope that we have addressed all the reviewers’ comments and look forward to your response.

Kind regards
Tracy Kolbe-Alexander
Detailed response to reviewers:

**Reviewer 1:**

**Introduction:**

1. *The intro would benefit from at least some discussion of the morbidity/mortality attributed to NCD risk factors in South Africa (the most recent global burden of disease study seems an obvious reference for this) in order to highlight their importance as a public health problem in the country.*

We have included information relating to the morbidity and mortality attributed to NCD risk factors using the recent Global Burden of Disease publication by Lim et al., 2012, Lancet. The following has been added to the introduction on page 3:

“Furthermore, the burden of disease due to NCD risk factors was higher in 2010 than in 1990, and physical activity together with poor dietary habits accounted for 10% of global Disability life years (DALYS) in 2010 [5]. Low fruit and vegetable intake and physical inactivity each accounted for 1.1% of total DALYs and were ranked as the 11th and 12th leading risk factors, respectively, in South Africa [6].”

2. *The authors state that describing the clustering of risk factors among those who volunteer for the HRA would help inform strategies to recruit new participants. How, exactly, would this do so? This should be made clear to the reader.*

The following text has been added on page 4 in the Introduction.

“Indeed, previous research has shown that employees completing HRA’s are usually older, report fewer days of sick leave and have better self-reported health status, than those who do not [13, 14]. … Thus the messaging and advertising encouraging HRA participation could be directed at encouraging younger and less healthy employees to participate. In addition, some employees have listed lack of time as a barrier for participation, thus future strategies might include having shorter HRA’s in order to encourage higher rates of participation [14].”

3. *The intro lacks any discussion of the insurance context of SA. The stated aim is to look at clustering and expenditures among privately insured employees – are they the majority? What does this insurance cover? How do they compare with publicly insured individuals? Without this information it is unclear as to how the findings will inform policy.*

We agree with the reviewer that information relating to private health insurance in South Africa was missing from the introduction. We have added the following to the ‘Introduction’ on page 4.

Consequently, some private health insurance companies offer wellness days and opportunities for employees to complete a HRA in order to determine their current health status. Private health insurance coverage is relatively low in South Africa where only 16% of South Africans have private health insurance [16] [17]. Being
employed has been identified as one of the main predictors of having private healthcare insurance [18], which might be largely due to some worksites including compulsory private healthcare cover as a condition of service [19].

Methods:
1. **There are potential biases introduced by this being a wellness day that is sponsored by either the insurer or the employer (it’s not clear from the manuscript who is actually paying/hosting this event and whether their role is made explicit to the employees). These need to be addressed.**

We agree with the reviewer that there are potential biases that can be introduced, and we address this in the discussion on page 11 and the ‘strengths and limitations’ on page 14. However, we have also included addition information in the “Methods” section of the paper on page 5, to try to further address the reviewer’s concerns.

Accredited service providers who conduct all the measurements at the wellness days are required to comply to the health insurer’s standard measurement guide and calibration of equipment in order to ensure that all measures are accurate and reliable. The cost of the wellness day is shared between the private health insurer and the employer, thus the employee does not make any financial contribution towards the services.

2. **Data on physical activity were collected over a 12 month period. Should any seasonal effects be considered?**

The reviewer is correct that there is a potential for seasonal effects on habitual levels of physical activity. However, employees reported on the amount of physical activity in a ‘usual’ week, and not ‘the last seven days’ in an attempt to reduce this effect.

3. **There are many potential issues with the participants that need to be addressed. They were wellness participants from 68 companies and their expenditure was compared to non-participants. What were these companies?**

The data was obtained from all the companies who had wellness days over a 12-month period. These companies were from various sectors including engineering, logistics, consultants, information technology, academic, financial and transport sectors. This has been included in the methodology on page 5.

These worksites included companies from various sectors including engineering, logistics, consultants, information technology, academic, financial and transport sectors.

Were the non-participants from the same company? If not, is there something different about the nature of their occupation that might affect their physical fitness and their expenditure? If there is anything specific to the companies that might affect any of the outcomes, clustering might need to be accounted for (e.g. with a multilevel model).

We agree with the reviewer that the type of company / sector could affect physical
fitness and healthcare expenditure. However, the non-participants were based at the same companies as those who chose to attend, thus addressing the influence of the type of company on outcomes. In addition, although there were different sectors represented, most of the employees were office based.

4. The authors state that demographic information was collected. What demographic information? There are so many potential factors that can affect health expenditure (gender, age, occupation, education, socio-economic status, marital status...). Data on these would need to be collected and controlled for but no details on this are given in the manuscript.

The demographic information included gender and age and this has been included in the methods on page 6. These variables were included in the multiple linear regression described in the methods section. The following has been added to the text on page 9;

“...The variables included in this model were age, use of chronic medication and number of total number of modifiable risk factors (excluding physical activity) for NCD. Likewise, a multiple regression analysis for doctors visits was performed which included age, use of chronic medication and modifiable risk factors (excluding physical activity) and habitual levels of physical activity.”

5. In reference to measurement, the authors state that self-reported physical activity was used. The reader would need to be convinced (i.e. with reference to literature) that this is a reliable measure (e.g. do people tend to over-report their physical activity?)

Although self-reported measures of physical activity have some limitations, it has been widely used due to their decreased participant burden and being cheaper than objective measures. This method of data collection for physical activity has therefore been regarded as acceptable in the literature (Dishman et al., 2001; Prince et al., 2008). In a cross-sectional sample of just over 13000 members completing HRA in 2009, there was a significant correlation between self reported PA with gym visits (rho= 0.4278, P<0.0001) (Nglazi, in preparation).

We have tried to address the reviewer’s comment, by including the following in the ‘Limitations’ section of the manuscript on pages 14 and 15;

Self-reported physical activity could be viewed as a limitation for this research study, as objective measures might be able to provide more accurate data. However, self-reported measures of physical activity have been generally accepted among researchers due to their lower cost (than objective measures) and feasibility when including larger numbers of participants [32] [33].

6. A sentence could be added about who the 'trained staff' are who are collecting the clinical measurements.

The following has been added to the text on page 6 in the ‘Methods’ section: Trained staff, including Exercise Physiologists and nurses, performed...
7. **The statistical analysis needs major revision. Why was only a one-way ANOVA conducted? Why weren't any very relevant confounders (age, etc.) adjusted for (especially as it seems data were collected on these)? Had the authors considered using a matching approach? Also with the test of differences in risk factors between those who met physical activity guidelines and those who did not - there are so many confounders to consider and this is not addressed.**

The ANOVA was used in the statistical analysis to determine if there were any significant differences in healthcare expenditure between attendees and non-attendees, in addition to try to determine whether those who met the physical activity guidelines had lower healthcare expenditure, than those who did not. Age and gender were entered as covariates in this model. We have clarified this in the methods section, under the 'statistical analysis' sub-heading.

Furthermore, we aimed to address the issue of confounders by conducting the multiple linear regression which was used to determine factors associated with meeting the physical activity guidelines. This has been clarified in the methodology, under the heading, 'Statistical Analysis' on page 9;

"An analysis of variance (ANOVA) was used to determine if there were differences in healthcare expenditure between persons meeting physical activity guidelines versus those who were insufficiently active. Similarly, participants were grouped into those with more than 2 risk factors for CVD and those with less than two risk factors, and an ANOVA was performed to determine whether there were significant differences in healthcare expenditure between these two groups. Age and gender were used as covariates in both ANOVA models.

Multiple linear regression analysis was conducted to examine the factors associated with meeting physical activity guidelines. The variables included in this model were age, use of chronic medication and number of total number of modifiable risk factors (excluding physical activity) for NCD. Likewise, a multiple regression analysis for doctors visits was performed which included age, use of chronic medication and modifiable risk factors (excluding physical activity) and habitual levels of physical activity."

**Results:**

1. **The authors address logical constraints that caused non-participation such as being 'off-site' but what about other causes, such as ill health, that might actually affect the results?**

We agree with the reviewer that being absent from work due to illness and thus missing the wellness day might affect results, and this is addressed in the first paragraph of the discussion. We have included being absent from work due to illness in the text on page 9:

The non-participants were comprised of those who chose not to attend the wellness days, **who were not at work due to illness**, or those who were unable to attend due to logistical constraints, such as being off-site or on another shift.
2. The authors introduce the use of a regression model in the results section - this should be in the methods section. Moreover, the covariates controlled for in the model (if any??) should be identified. Lastly, it is used to show that physical activity results in increased odds of overweight or obesity, which is something that is known. Why not use it for the other (more interesting?) question on expenditure?

The multiple regression is introduced in the Methodology section under the heading, 'Statistical Analysis'. We have tried to address the reviewer’s comment by adding more detail and having a separate paragraph for the multiple regression component of the analysis. The following text has been added on pages 9;

“Multiple linear regression analysis was conducted to examine the factors associated with meeting physical activity guidelines. The variables included in this model were age, use of chronic medication and number of total number of modifiable risk factors (excluding physical activity) for NCD. Likewise, a multiple regression analysis for doctors visits was performed which included age, use of chronic medication and modifiable risk factors (excluding physical activity) and habitual levels of physical activity.

Finally logistic regression models were computed to determine the odds of being classified as ‘at risk’ for each of the other NCD risk factors in addition to being insufficiently active. Additional logistic regression analyses were performed to determine the odds of hospitalization and an additional visit to the doctor based on each year that the ‘Vitality Risk Age” is more than chronological age.”

We agree with the reviewer that the relationship between physical activity and obesity is known. The results sub-section on page 10 aims to determine the relationship between physical activity and the risk factors for NCD, as this was one of the aims of the paper. As per the reviewer’s comment, we also explore the relationship between physical activity and healthcare expenditure, and this is reported on page 11, paragraph 2.

Minor Compulsory Revisions

1. In para 3 of the introduction, the word ‘who’ is missing (further insight into the characteristics of individuals WHO choose...)

This has been corrected and ‘who’ has been added.

2. The authors sometimes refer to the insurer as a ‘private insurer’ and sometimes as the ‘national health insurer’. This would be confusing to those who are unfamiliar with the South African insurance context and who might assume ‘national’ means ‘public’.

We would like to thank the reviewer for this comment and have ensured consistency by using ‘private insurer’ as suggested by the reviewer.
Reviewer 2:

Methods
1. The methods section implies that the ‘annual wellness days’ are conducted annually. For participants who participated, it is important to know whether the data used refers to the most recent participation; aren’t there some who participated multiple times? If so, which occasion was used?

The wellness days are conducted once per year at companies, therefore employees cannot participate multiple times per year. We have clarified this in the text on page 5; “Each worksite only has one wellness day per year therefore the employees only have one opportunity to participate per calendar year. This analysis comprises an evaluation of data that were collected over a 12-month period (January to December), from 68 companies.”

2. Interesting why height was measured to ‘0.1’ cm

This error has been corrected to the nearest centimeter on page 7.

3. It is important to note in the methods that these data were extracted from records(i.e. secondary data), and that the different ‘wellness testing sites’ follow similar standard procedures; it would be also important to know about the integrity of the dataset – whether everyone for instance had all the parameters required in their medical record of the study selected only those with complete records.

We agree with the reviewer that it is important for each site to follow the same testing protocol, and this has been clarified in the text on page 5, by adding the following text;

“Accredited service providers who conduct all the measurements at the wellness days are required to comply to the health insurer’s standard measurement guide and calibration of equipment in order to ensure that all measures are accurate and reliable.”

4. Unless is was ANOVA for a linear regression, ANOVA for categorical variables is often used where there are more than 2 comparison groups; with 2 independent samples (numerical data) we use the t-test for data that fits the assumptions.

We would like to thank the reviewer for pointing out the error, which has been corrected in the text on page 9;

“An analysis of variance (ANOVA) was used to determine if there were differences in healthcare expenditure between persons meeting physical activity guidelines versus those who were insufficiently active. Similarly, participants were grouped into those with more than 2 risk factors for CVD and those with less than two risk factors, and an ANOVA was performed to determine whether there were significant differences in healthcare expenditure between these two groups. Age and gender were used as covariates in both ANOVA models.”
5. **A statement of ethical approval needed**

A statement of ethical approval is in the text on page 6 in the methodology: “Ethical approval for this research study was obtained from the University of Cape Town’s Research and Ethics Committee of the Faculty of Health Sciences (REF 348/2008).”

**Results**

6. **Under characteristics of participants: The statement that ‘men and women were of similar age’; maybe you meant the different between means is not statistically significant.**

The reviewer is correct that the difference in mean age for men and women is not statistically significant. The text has been revised as recommended on page 10;

“The mean age for men and women was not statistically different (Table 2).”

7. **The second statement is also erroneous ‘placing them in the overweight category’ when you are referring to the mean; instead refer to the ‘mean’ not ‘them’ because many of them are not.**

We would like to thank the reviewer for this comment and have revised the text on page 10;

“The mean Body Mass Index (BMI) was more than 25 for both men and women, which is in the overweight category.”

8. **Under ‘Health Expenditure’ it brings unnecessary confusion if you say those who met the guidelines were classified as ‘not at risk’ – why not simply refer to them as they are ‘met the guidelines’.**

The following sentence has been deleted from the ‘Health Expenditure’ section; Employees not meeting the physical activity recommendations were classified as being ‘at risk’ while those meeting guidelines were ‘not at risk’.

In addition, the text has been revised on page 11;

“The multiple regression model which included age and use of chronic medication, showed that meeting physical activity guidelines had a marginal but significant association with the number of visits to the doctor ($r^2 = 0.04$; p<0.001).”

9. **The multiple regression model you refer to in this sub-section and therefore the r-square – is it a multiple linear regression model or a logistic regression model. The way it is stated, it is difficult to know what you are comparing with what.**

These results are based on the multiple linear regression model. We addressed the reviewer’s comment by revising the text on page 11;
“The multiple regression model which included age and use of chronic medication, showed that meeting physical activity guidelines had a marginal but significant association with the number of visits to the doctor (r² = 0.04; p<0.001). Those who were insufficiently active had significantly higher number of visits than those who were meeting physical activity guidelines, 2.91 (95% CI: 2.8; 3.0) and 2.67 (95% CI: 2.44; 2.88), respectively, p=0.04. Similarly, employees with two or more risk factors had significantly higher numbers of visits to the doctor in a 12-month period, and this was coupled with significantly higher health-care expenditure (related to doctors’ visits), after adjusting for age and gender (Table 5).”

In addition, the following has been added to the text in the methods section, under the ‘statistical analysis’ section on page 9;
“Finally logistic regression models were computed to determine the odds of being classified as ‘at risk’ for each of the other NCD risk factors in addition to being insufficiently active. Additional logistic regression analyses were performed to determine the odds of hospitalization and an additional visit to the doctor based on each year that the ‘Vitality Risk Age” is more than chronological age.”

10. Provide statistical evidence for the following assertions where possible:
   • Para2, sentence 1 (under the subtitle wellness participants vs non-participants)

   Only participants who are registered for chronic medication are able to incur costs for the chronic medication listed in Table 1. However, we aimed to address the reviewer’s comment by changing the sentence on page 10 to;

   “Wellness day participants had significantly lower chronic medication-related costs than the non-participants (Table 1).”

   • Under participant characteristics: for the difference between age of women and men

   The age for the men and women, together with the p-value for the one-way ANOVA is presented in Table 2 and the text on page 9 directs the reader to Table 2. In addition, we have included the following text in the ‘Statistical Analysis’ section on page 8;

   “In addition, the One-Way ANOVA was computed to determine if there were any significant differences between the men and women who participated in the wellness day.”

   • Under health care expenditure: Para 3, the statement on employees with more than 2 risk factors having more visits to facilities

   This data is presented in the new Table 5 (based on Reviewer 3’s recommendations).
11. **Affirming what was noted in the methods, no where do we see ANOVA results (except for where you quote an r-square and even this is not clear).**

We hope that we have addressed this comment in the previous comment where we clarify that we use the one way ANOVA.

**Minor essential revisions**

**Introduction**
The introduction is sufficiently informative. There are minor language issues though.

12. **Paragraph 1, line 5: It is better to use ‘low physical activity’ or ‘insufficient physical activity’, or ‘sedentariness’ instead of inactivity.**

The text has been corrected to replace inactivity with ‘insufficient physical activity’.

13. **Paragraph 3, second last line ‘individuals ‘who’ choose’...**

The text has been corrected to add the word ‘who’.

14. **Last paragraph, sentence 2 is not complete the way it is written. Some key words in the object are missing.**

We are a little unsure of which words in the object are missing from this sentence. However, we tried to correct this sentence and it reads as follows:

Therefo*re, the main aim of this study was to determine the extent to which insufficient physical activity clustered with other risk factors for NCD in employed persons presenting for health risk assessment as part of worksite wellness day.*

**Tables**

15. **Table 1: Instead of ‘participants’ vs. ‘non-participants’ can you use a phrase that is more clear e.g. ‘participated in wellness clinic’ vs. Did not**

The subheadings in Table 1 have been changed as recommended by the reviewer.

16. **Table 3: What do the figures in the table refer to? Are they frequencies?**

The figures in Table 3 refer to the number of employees with additional risk factors for NCD’s. We have changed the title of this table to:

“Physical activity risk and number of employees with additional risk factors for non-communicable disease”

**Discussions**

17. **Conclusions are not well stated – the study has very important findings but these are not**
well summarised e.g. majority of employees having more than 2 risk factors, majority of employees not reaching desired physical activity threshold and comparison of those who participated with those who did not. Also do not talk about odds in the conclusions – simply state excess risk in ordinary terms.

We have aimed to address the reviewer’s comment by revising the text in the conclusion on page 15 as follows;

“This research study has shown that more than two thirds of employees did not meet the recommended physical activity guideline of 150 minutes of moderate to vigorous intensity physical activity per week. These employees had a higher number of additional risk factors for NCDs compared to those who were sufficiently physically active. Moreover, those employees with two or more risk factors for NCD had significantly higher healthcare expenditure related to doctor’s visits than those with fewer risk factors.”
REVIEWER 3:

Major Revisions:

1. **The study title needs to convey clustering as major focus of the study as it is listed as main aim of the study.**

   The title has been changed to “Clustering of risk factors for non-communicable disease and healthcare expenditure in employees with private health insurance presenting for health risk appraisal: A cross-sectional study” as suggested by the reviewer.

2. **Authors speak about 7 risk behaviours and a score of 0 to 7 however there are only 6 risk behaviours as authors have added fruit and veg intake as one variable. This requires attention.**

   The reviewer is correct that we did not add a description of the 7th risk factor included in the calculation of total number of risk factors for NCD. We have corrected this error and included the following in the text on page 7;

   …..age (> 45 years for men and > 55 years for women)…

3. **There are two pathways by which PA affects chronic conditions: direct associations (some cancers and PA or CVD and PA) or through risk factors such as high BP, BMI or high cholesterol. Authors do not mention it at all. If the authors are combining these risk behaviours in one scale then this is important issue to add to the discussion.**

   We have added the following text to the discussion on page 13.

   “The association between physical activity and reduced burden of disease may be mediated by intermediates in the causal pathway. For example, Rennie et al., found that physical activity reduced the likelihood of metabolic syndrome for both men and women [28]. Moderate intensity physical activity has been inversely correlated to waist to hip ratio in women, and BMI, as well as total cholesterol and triglyceride concentrations in men [28]. Similarly, men and women participating in vigorous intensity physical activity were less likely to have metabolic syndrome, OR = 0.50; 95% CI 0.39, 0.64, even after adjusting for age and other NCD risk factors [28]. These findings are supported by Pronk and Kottke who reported that adults who are physically active have a more favourable bio-maker profile and lower rates of all cause mortality than those who are inactive [15]. Furthermore, a recent research study among university employees reported that those who were inactive reported greater interest in health promotion programs [29]. Thus both completing the HRA and the associated results might be a valuable teachable moment to improve lifestyle behaviors and health status.”

4. **Figure 1 is the only necessary figure and rest of the figures can be deleted as the data is included under results.**

   We have deleted figures 2, 3, 4 and 5 as suggested by the reviewer.
5. *Figure one does not explain origin of certain numbers: 409 mentioned in the abstract (either remove it or add it to the figure), 2789 from the abstract when it is not in Figure 1, 789 mentioned on page 5 (how does it fit in figure 1?).*

The n=409 has been removed from the abstract as suggested by the reviewer. Figure 1 has been corrected to include, n= 2789

The number 46789 on figure and in the text represents employees who are also members of the private health insurance but who did not participate in the wellness day.

6. *The number of participants included in each analysis needs to be mentioned clearly. Eg what was the number of participants that was use to calculate prevalence of employees at risk in figure two or first paragraph on page 9?*

The number of participants used to calculate prevalence of employees at risk has been included in the title of figure 2, and also in the text on page 10.

7. *One of the main aims is to test clustering of risk behaviours however the authors have missed on the literature that is available on clustering of health behaviours from the background section. The background needs to reduce information on HRA and missing information on clustering.*

As per the reviewer’s recommendation, the following text has been deleted from the introduction.

In addition to providing employees with individual health-related feedback, the HRA can also be used to establish future human capital expenditure that might be incurred by companies, as a consequence of health risk behaviors [8]. Thus, the HRA can provide the underlying rationale for, and the focus of targeted worksite health promotion programs in corporate settings [8].

We have included the following information in the introduction on page 3 as suggested by the reviewer;

“Physical inactivity appears to cluster with other risk factors for cardiovascular disease where those who are insufficiently physically active were more like to have additional risk factors such as elevated serum triglycerides, hypertension and elevate fasting glucose levels [7]. Similarly, the results of a study based on data from the United States National Health and Nutrition Examination Survey, shows that individuals who are inactive are more likely to have additional risk factors for cardio metabolic disease [8]. These findings are corroborated by a study conducted in Swedish men and women where those who had higher levels of physical activity had significantly lower triglycerides and less atherogenic lipid profiles compared to those who were inactive [9]. Furthermore, the Swedes who had higher fitness levels were 50% less likely to have additional three or more risk factors for NCD [9].”
8. Discussion requires mention of bias introduced due to large number of participants refusing to participate (46789), self-report data on physical activity, volunteer to participate.

The bias is discussed in the first paragraph of the discussion, and we have also added the following text on page 15:

“Another limitation of the research study is that a large number of employees did not participate in the wellness days. As a result, selection bias can occur whereby the healthier and more motivated employees attend the wellness days.”

9. In cross-sectional data it is unclear that physical activity is changed after the diagnosis of chronic condition or before, hence this need to be discussed as a part of discussion with mention of teachable moment in changing health behaviours.

We agree with the reviewer, and the following text has been added to the discussion on page 13.

“Furthermore, a recent research study among university employees reported that those who were inactive reported greater interest in health promotion programs [29]. Thus both completing the HRA and the associated results might be a valuable teachable moment to improve lifestyle behaviors and health status.”

10. Rather than adding a figure for odds ratio, a table would be better with OR, CI, P-values and comparators.

Figure 4 has been replaced with a Table as recommended by the reviewer.

11. A table for health care expenditure with groups (physical activity: at risk/not at risk), risk factors (less than 2/two or more) : with numbers, mean expenditure with (CI) and p values will add value.

This table has been added as recommended by the reviewer.

Minor

1. Paragraph 2 on page 11 requires precise discussion and “the ability----overall health status” is unnecessary and does not add value to the paper.

This paragraph has been deleted as recommended by the reviewer.

2. Paragraph two on page 12 starts with explaining use of health services by participants with higher risk score and then moves on to talking about PA. It is unclear, what point do authors want to convey.

We have addressed the reviewer’s comment by revising the text on page 14 as follows:

“Since nearly three-quarters (71%) of the participants in our study were insufficiently physically active, most of the employees with more than two risk factors for NCD were inactive. Therefore, our findings are in agreement with previous research that reported habitual levels of physical activity were inversely associated with health care expenditure [30].”