PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Factors associated with self-reported health among New Zealand military veterans: a cross-sectional study.
AUTHORS	McBride, David; Samaranayaka, Ariyapala; Richardson, Amy; Gardner, D; Shepherd, Daniel; Wyeth, Emma; de Graaf, Brandon; Derrett, Sarah

VERSION 1 – REVIEW

REVIEWER	Porter, Ben Mississippi State University
REVIEW RETURNED	29-Sep-2021

GENERAL COMMENTS	The manuscript "Factors associated with self-reported health among New Zealand military Veterans: A cross-sectional study" presents results from an online sample of 1,817 Veterans collected from a variety of different sources. The study has several strong benefits including the sample itself and the variety of sources used to attract participants. However, there were several issues that I noticed while going through the manuscript that I thought merited attention. I have listed these below separated out by their significance. Major Issues 1) The coding for sex seems to be off based on the values of the univariate table. That beta should be ~3.9 based on the supplementary table. It looks like .65 might be the difference between men and missing categories 2) I feel like the end points of VAS (best and worst day imaginable) indicate that the measure itself is relative health. So higher scores on this should only be influenced by state level factors. By predicting this, you aren't really saying what makes health better but what factors are associated with feeling healthier on one day than another. 3) The analysis you are doing is more pairwise deletion than listwise deletion. The changing sample size makes it difficult to compare nested models (e.g., models 1 and 2). I would keep with the N=1,557 for this section at the very least but would also recommend using this sample for the entire paper so all the analyses use the same sample. 4) The AAQ-II seems like it would correlate highly with PTSD (e.g., emotional stuck points, etc.) With this being a military sample, PTSD is probably more prevalent than in the general population. If it is available, it would probably make sense to covary for it. 5) I am not following the conclusions regarding the application of ACT. Are the authors suggesting that ACT should be used for individuals experience poor physical health? Minor Issues

	1) It seems like the definition of Veteran being used in the current paper is "having been deployed to a particular conflict" rather than "having served in the military but is no longer serving" I think this is likely a difference from my own background working with the US military, but would it be possible to clarify this? 2) On line 131, I am not familiar with Action and Commitment Therapy. Is this a mistype for Acceptance and Commitment Therapy? 3) More information needs to be included on the scales themselves 4) Include z-tests for the differences in any domain. From the graph provided, it looks like self-care is also significantly different, albeit a smaller difference. 5) For the issue with social support, looking at a correlation matrix would help determine what factors are associated with it which may be responsible to attenuating the association. 6) I think that it is important to point out that the factors in the multivariate regression are unique predictors of the outcome. Overall associations are shown in Table 1, but this distinction isn't clear from the discussion. 7) The section "Comparisons with other studies" seems like it would fit better in the introduction or would need to be compared and contrasted to the current study. 8) The brevity of the scales used should be listed as a limitation as reliability is not calculable.

REVIEWER	Maule, Alexis Army Public Health Center
REVIEW RETURNED	10-Oct-2021

GENERAL COMMENTS	Overall, a well written study. However, I believe that the collinearity of several of the independent variables should be examined and discussed in the manuscript. And a discussion regarding the study sample and potential differences that exist between military personnel currently serving (the vast majority of your sample) and veterans who are no longer serving. There can be major differences in the available resources and health care access between these two groups. I see this as the largest limitation in your examination of "veteran" health.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1. Dr. Ben Porter, Mississippi State University Comments to the Author:
The manuscript "Factors associated with self-reported health among New Zealand military Veterans: A cross-sectional study" presents results from an online sample of 1,817 Veterans collected from a variety of different sources. The study has several strong benefits including the sample itself and the variety of sources used to attract participants. However, there were several issues that I noticed while going through the manuscript that I thought merited attention. I have listed these below separated out by their significance.

Major Issues
1) The coding for sex seems to be off based on the values of the univariate table. That beta should be ~3.9 based on the supplementary table. It looks like .65 might be the difference between men and missing categories

Thanks to the reviewer for detecting this typo. Univariable table is correct, but row labels were swapped in the supplementary table. It is corrected now as below.

Gender N mean SD
female 220 75.0 16.9
male 1520 74.4 17.2
missing 27 70.3 17.2

2) I feel like the end points of VAS (best and worst day imaginable) indicate that the measure itself is relative health. So higher scores on this should only be influenced by state level factors. By predicting this, you aren't really saying what makes health better but what factors are associated with feeling healthier on one day than another.

The EQ-5D-5L is one of the most widely used health related quality of life measures. In response to the reviewer's comment, the EQ-5D-5L's VAS scale does not ask about the best and worst 'day' imaginable; i.e. people are not being asked to rate their health today relative to their health on other days. Instead, it asks people to rate their health 'today' on a scale ranging from 100 (best health imaginable) to 0 (worst health imaginable).

3) The analysis you are doing is more pairwise deletion than listwise deletion. The changing sample size makes it difficult to compare nested models (e.g., models 1 and 2). I would keep with the N=1,557 for this section at the very least but would also recommend using this sample for the entire paper so all the analyses use the same sample.

We agree with the reviewer that in the univariable component of the analysis was by pairwise deletion. We chose that to maximise the use of available information in the data. We also agree that the results of models 1 and 2 in table 2 are difficult to compare due to different numbers of participants. However, this was not the intention: As mentioned in the statistical analysis section, the intent was to identify variables associated with the outcome (as opposed to estimating independent effects of variables that were already identified). Therefore the model prior to variable selection is not the final model, it is an intermediate step but presented here to provide a holistic view of the adjusted results, and our conclusions are based on the final model, based on listwise deletion of variables.

4) The AAQ-II seems like it would correlate highly with PTSD (e.g., emotional stuck points, etc.) With this being a military sample, PTSD is probably more prevalent than in the general population. If it is available, it would probably make sense to covary for it.

We haven't looked at PTSD in this paper, but we have examined PTSD as an outcome in a separate paper published in PLOS ONE; it is cited as Reference 26. https://doi.org/10.1371/journal.pone.0231460.

5) I am not following the conclusions regarding the application of ACT. Are the authors suggesting that ACT should be used for individuals experience poor physical health?

Act has been shown to improve pain acceptance, we have added a new reference by Hughes et al. (35).

Minor Issues
1) It seems like the definition of Veteran being used in the current paper is "having been deployed to a particular conflict" rather than "having served in the military but is no longer serving" I think this is likely a difference from my own background working with the US military, but would it be possible to clarify this?

Thank you, this tends to cause confusion. In working with Veterans, we prefer to use the more generally accepted definition, i.e. having served in the military. In New Zealand, we have two classes of Veterans, legal Veterans as defined by the Veteran Support Act, who get support from Veterans affairs, (line 60 et seq.) The Ministerial panel (line 68) funded the study, and specifically charged us to look at this group. We did not however exclude any Veteran from the study, but controlled for this in the analysis.

2) On line 131, I am not familiar with Action and Commitment Therapy. Is this a mistype for Acceptance and Commitment Therapy?

Thank you, this has been corrected

3) More information needs to be included on the scales themselves

Basic information on each scale was given in the questionnaire section, but have not included measures of validity and internal consistency we are writing for a general audience. We do acknowledge this in the 'limitations' section, line 232 et seq.

4) Include z-tests for the differences in any domain. From the graph provided, it looks like self-care is also significantly different, albeit a smaller difference.

Thank you, we have now done Z-tests to compare the proportion of veterans with 'any problem' in each EQ-5D dimension with the corresponding proportion in the NZ general population. We find that the proportion with 'any problems' is significantly higher among veterans in each dimension except the anxiety or depression dimension. This information is now added to the results section as table 1 and included in the first sentence of the discussion.

5) For the issue with social support, looking at a correlation matrix would help determine what factors are associated with it which may be responsible to attenuating the association.

Thank you for this insight. The correlation matrix shows a significant correlation between social support score and GHQ12 score. This could be a possible reason for not retaining social support score in the final model after adjusting for GHQ12 as mentioned in 2nd para of discussion, and commented upon in the discussion, line 232.

6) I think that it is important to point out that the factors in the multivariate regression are unique predictors of the outcome. Overall associations are shown in Table 1, but this distinction isn't clear from the discussion.

Thank you, we now make this distinction clear in discussion, line 204.

7) The section "Comparisons with other studies" seems like it would fit better in the introduction or would need to be compared and contrasted to the current study.

We agree, and now 'compare and contrast,' line 257 et seq.

8) The brevity of the scales used should be listed as a limitation as reliability is not calculable.

We now do this, line 232.

Reviewer: 2. Dr. Alexis Maule, Boston University School of Public Health, US Army Research Institute of Environmental Medicine Comments to the Author:

Overall, a well written study. However, I believe that the collinearity of several of the independent variables should be examined and discussed in the manuscript. And a discussion regarding the study sample and potential differences that exist between military personnel currently serving (the vast majority of your sample) and veterans who are no longer serving. There can be major differences in the available resources and health care access between these two groups. I see this as the largest limitation in your examination of "veteran" health.

We agree, and collinearity was evident with social support, which was dropped. We discuss this in line 208 et seq.

In practice serving personnel have good access to primary care, however if tertiary level care is required for illness, they get this through the public system and everyone is covered for accidental injury. To get 'cover' from Veterans affairs, the injury must be associated with a particular deployment. We explain this in line 222 et seq.

VERSION 2 – REVIEW

REVIEWER	Porter, Ben Mississippi State University
REVIEW RETURNED	28-Feb-2022

GENERAL COMMENTS	Thank you for responding to several of my comments. I still have major concerns about this manuscript that were not addressed in the responses.
	3) The analysis you are doing is more pairwise deletion than listwise deletion. The changing sample size makes it difficult to compare nested models (e.g., models 1 and 2). I would keep with the N=1,557 for this section at the very least but would also recommend using this sample for the entire paper so all the analyses use the same sample.
	We agree with the reviewer that in the univariable component of the analysis was by pairwise deletion. We chose that to maximise the use of available information in the data. We also agree that the results of models 1 and 2 in table 2 are difficult to compare due to different numbers of participants. However, this was not the intention: As mentioned in the statistical analysis section, the intent was to identify variables associated with the outcome (as opposed to estimating independent effects of variables that were already identified). Therefore the model prior to variable selection is not the final model, it is an intermediate step but presented here to provide a holistic view of the adjusted results, and our conclusions are based on the final model, based on listwise deletion of variables.
	Using pairwise deletion can lead to incorrect assumptions. If you want to use all data available, you should use either FIML or multiple imputation. Barring that, a consistent sample should be used throughout the manuscript (I've been guilty of this in the past, but it really is best practices). If you aren't interested in the intermediate model, I would only present the final model particularly as the difference is trimming nonsignificant effects. Alternatively, if you wanted to show which variables are independently associated with the outcome, I would stick with the adjusted model.
	4) The AAQ-II seems like it would correlate highly with PTSD (e.g.,

emotional stuck points, etc.) With this being a military sample, PTSD is probably more prevalent than in the general population. If it is available, it would probably make sense to covary for it.

> We haven't looked at PTSD in this paper, but we have examined PTSD as an outcome in a separate paper published in PLOS ONE; it is cited as Reference 26. https://doi.org/10.1371/journal.pone.0231460.

I think you should covary for PTS or PTSD in this analysis given the data are available.

5) I am not following the conclusions regarding the application of ACT. Are the authors suggesting that ACT should be used for individuals experience poor physical health?

> Act has been shown to improve pain acceptance, we have added a new reference by Hughes et al. (35).

I think I might be unclear from the paragraph, but I don't know feel like the results are so strong to draw the conclusion from this study that ACT should be used to treat poor physical health among veterans. You might just want to relax the language surrounding this.

Minor Issues

1) It seems like the definition of Veteran being used in the current paper is "having been deployed to a particular conflict" rather than "having served in the military but is no longer serving" I think this is likely a difference from my own background working with the US military, but would it be possible to clarify this?

> Thank you, this tends to cause confusion. In working with Veterans, we prefer to use the more generally accepted definition, i.e. having served in the military. In New Zealand, we have two classes of Veterans, legal Veterans as defined by the Veteran Support Act, who get support from Veterans affairs, (line 60 et seq.) The Ministerial panel (line 68) funded the study, and specifically charged us to look at this group. We did not however exclude any Veteran from the study, but controlled for this in the analysis.

I appreciate the increased information in the manuscript. I still think I am unclear. Is the current study derived from currently serving personnel, personnel from prior conflicts (pre 1991 Gulf War), and individuals who found the survey?

3) More information needs to be included on the scales themselves

> Basic information on each scale was given in the questionnaire section, but have not included measures of validity and internal consistency we are writing for a general audience. We do acknowledge this in the limitations' section, line 232 et seq.

	I believe BMJ Open is intended for a scientific community. I would still include these metrics when they are able to be calculated. It's not necessarily an issue if internal consistency is low, but that would be expected to attenuate associations and potentially be responsible for the non-significant associations.

REVIEWER	Maule, Alexis Army Public Health Center
REVIEW RETURNED	04-Mar-2022

GENERAL COMMENTS	Authors have sufficiently address my previous comments on the manuscript, and also appear to have addressed those of the other reviewer. I have no further comments on the manuscript.

VERSION 2 – AUTHOR RESPONSE

Thank you for responding to several of my comments. I still have major concerns about this manuscript that were not addressed in the responses.

3) The analysis you are doing is more pairwise deletion than listwise deletion. The changing sample size makes it difficult to compare nested models (e.g., models 1 and 2). I would keep with the N=1,557 for this section at the very least but would also recommend using this sample for the entire paper so all the analyses use the same sample.

Now done, no pairwise deletions.

Using pairwise deletion can lead to incorrect assumptions. If you want to use all data available, you should use either FIML or multiple imputation. Barring that, a consistent sample should be used throughout the manuscript (I've been guilty of this in the past, but it really is best practices). If you aren't interested in the intermediate model, I would only present the final model particularly as the difference is trimming nonsignificant effects. Alternatively, if you wanted to show which variables are independently associated with the outcome, I would stick with the adjusted model.

AS has now kindly re-analysed the data as suggested, with new Tables 1-3, and have taken the 'consistent sample' option so that both models use the same participants.

4) The AAQ-II seems like it would correlate highly with PTSD (e.g., emotional stuck points, etc.) With this being a military sample, PTSD is probably more prevalent than in the general population. If it is available, it would probably make sense to covary for it.

I think you should covary for PTS or PTSD in this analysis given the data are available.

Thank you for this suggestion, the PCL-M has been added as an explanatory variable.

5) I am not following the conclusions regarding the application of ACT. Are the authors suggesting that ACT should be used for individuals experience poor physical health?

I think I might be unclear from the paragraph, but I don't know feel like the results are so strong to draw the conclusion from this study that ACT should be used to treat poor physical health among veterans. You might just want to relax the language surrounding this.

Thanks for clarifying this, that was not our intention. As per your suggestion, we agree that the results do not have sufficient weight to make this assertion, so the references were removed.

Minor Issues
1) It seems like the definition of Veteran being used in the current paper is "having been deployed to a particular conflict" rather than "having served in the military but is no longer serving" I think this is likely a difference from my own background working with the US military, but would it be possible to clarify this
It is difficult with our 'class system' of Veterans, this has, hopefully been clarified in lines 90-93.
3) More information needs to be included on the scales themselves

I believe BMJ Open is intended for a scientific community. I would still include these metrics when they are able to be calculated. It's not necessarily an issue if internal consistency is low, but that would be expected to attenuate associations and potentially be responsible for the non-significant associations.

We have now calculated Cronbach alpha, and this information is now included in table 2. We have reasonable levels of internal consistency, except perhaps for the audit C.