

## 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.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Self-reported physical and mental health of Australian carers: A cross-sectional study
<b>AUTHORS</b>	Hussain, Rafat; Wark, Stuart; Dillon, Gina; Ryan, Peta

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Professor Sally Chan University of Newcastle, Australia
<b>REVIEW RETURNED</b>	26-Mar-2016

<b>GENERAL COMMENTS</b>	<p>The aim of the study needs to be clarified. If this study aimed to establish a profile of self-reported health in rurally-based informal carers, the comparisons to Australian population norm have to be removed from the Results section. The authors may discuss this in the Discussion section but not present in Results.</p> <p>There are well-established evidences that many informal carers of people with chronic health conditions have poor physical and mental health status. It is also known that younger caregivers experienced more stress than older caregivers because of various responsibilities in life. The study presented little new knowledge on the health status of informal caregivers.</p> <p>The authors may want to compare their results with previous studies that were conducted in metropolitan areas so that the unique health problems of this target group may be identified.</p>
-------------------------	---

<b>REVIEWER</b>	Michele Peters University of Oxford, UK
<b>REVIEW RETURNED</b>	02-May-2016

<b>GENERAL COMMENTS</b>	<p>This is a manuscript on the physical and mental health of Australian carers. This is an important topic for research, as it is important to assess outcomes in carers but also as this focuses on a population that is likely to be less researched. Overall the article is very clearly written and well presented.</p> <p>The main comment is that the results of factors associated with self-reported health is a little confusing as many individual factors are discussed, possibly also as it includes a discussion of the rural carer scores to normative carer scores (the latter may be better placed in the discussion). It would be good to pull all these findings together in a regression analysis. Additionally, the SF-36 scores are not presented for all carer related variables e.g. the condition being cared for, relationship to the person being cared for etc. This would</p>
-------------------------	---

	<p>be an interesting addition the the paper. As work is not the significant factor for worse outcomes in younger carers, could it be that it is the relationship to the carer (e.g. caring for a parent is more stressful or burdensome than caring for a spouse, possibly related to living or not living in the same household). The authors could focus on the significant relationships only and present additional material in a supplement?</p> <p>A minor comment is that there is no key provided for the superscript letters in Table 2. Presumably these refer to the levels of significance?</p> <p>Out of interest, is there any other SF-36 data from Australian carers? It would be of interest to know if outcomes are different between rural and urban carers.</p> <p>Finally, it would be helpful if the authors could comment on potential self-selection bias in the limitations.</p>
--	---

### VERSION 1 – AUTHOR RESPONSE

#### Reviewer One

Comment 1: The aim of the study needs to be clarified. If this study aimed to establish a profile of self-reported health in rurally-based informal carers, the comparisons to Australian population norm have to be removed from the Results section. The authors may discuss this in the Discussion section but not present in Results.

#### Authors' Response

Part A: Clarity in Aim – Agree with comment.

Action: Manuscript text revised in response to the reviewer's comments [see p. 5]

Part B: Comparison with Australian norms - Both reviewers have commented on this issue. On review we agree with the reviewers' comments.

Action:

1. Manuscript revised to remove discussion of Australian normative data for outcome variables from Results section. Comparison of findings from present study and Australian normative data has been moved to Discussion section and appropriately revised to link in better with discussion of other relevant issues.
2. Figure 2 has been revised to only reflect study data.
3. Figure 3 (original Figure 2) has been placed in supplementary materials section. We feel it provides useful comparative information but leave the final decision of its inclusion to the Editor.

Comment 2: There are well-established evidences that many informal carers of people with chronic health conditions have poor physical and mental health status. It is also known that younger caregivers experienced more stress than older caregivers because of various responsibilities in life. The study presented little new knowledge on the health status of informal caregivers.

#### Authors' Response

Evidence from other studies

Action:

1. We have included more information on carer studies from Australia; see revised manuscript, p.4 (Introduction) and pp.15-17 (Discussion).
2. We undertook another exhaustive review of PubMed, Medline, ProQuest and other relevant

databases, and are unable to locate any Australian study that included an urban and rural comparison on carer health.

We disagree with the esteemed reviewer about generalizability of studies from urban areas of Australia to rural regions. If this was the case, there would not be such a gap which has been acknowledged by all national bodies including Seniors Australia, Carers Australia and the Commonwealth government itself.

All authors have considerable work experience of working in the Australian rural health sector and academia (combined over 50 years). Issues in carer health in rural areas have not been assessed in a systematic manner beyond some qualitative studies. There is no Rural Carer study that has used comprehensive and internationally validated outcome measures such as SF-36, CES-D and K-10 concurrently. The Australian Bureau of Statistics survey of Disability, Ageing & Carers Survey contains no validated outcome measures [only self-reported individual health issues in a somewhat inconsistent manner]. The ABS-managed Household Labour & Income Dynamics survey [HILDA] has a small rural sampling frame but no explicit focus on caregiving. Therefore, we feel the current state-wide study provides information that is useful for understanding the health issues of rural carers.

#### Reviewer Two

Comment 1: This is a manuscript on the physical and mental health of Australian carers. This is an important topic for research, as it is important to assess outcomes in carers but also as this focuses on a population that is likely to be less researched. Overall the article is very clearly written and well presented.

#### Authors' Response

No action item requested.

Comment 2: The main comment is that the results of factors associated with self-reported health is a little confusing as many individual factors are discussed, possibly also as it includes a discussion of the rural carer scores to normative carer scores (the latter may be better placed in the discussion).

#### Authors' Response

Agree with the comment.

Action:

1. Manuscript revised to remove discussion of Australian normative data for outcome variables from Results section. Comparison of findings from present study and Australian normative data has been moved to Discussion section and appropriately revised to link in better with discussion of other relevant issues.
2. Figure 2 has been revised to only reflect study data.
3. Figure 3 (original Figure 2) has been placed in supplementary materials section. We feel it provides useful comparative information but leave the final decision of its inclusion to the Editor.

Comment 3: It would be good to pull all these findings together in a regression analysis.

#### Authors' Response

Agree with the comment.

Action: We have undertaken regression analyses for all four outcome variables. The techniques have been discussed in the Methods section and the results are presented both in the Results section and in an additional table (see Table 3).

Comment 4: Additionally, the SF-36 scores are not presented for all carer related variables e.g. the

condition being cared for, relationship to the person being cared for etc. This would be an interesting addition the paper.

As work is not the significant factor for worse outcomes in younger carers, could it be that it is the relationship to the carer (e.g. caring for a parent is more stressful or burdensome than caring for a spouse, possibly related to living or not living in the same household). The authors could focus on the significant relationships only and present additional material in a supplement?

**Authors' Response**

Agree with the comment.

Action: We have provided results for additional variables in an expanded and enhanced Table 2. We have also simplified the notations for significant results for Table 2.

The difficulty in presenting additional material in a supplement is that segmentation of data analyses and representation is not clear cut and we would need to almost repeat a considerable component of Table 2 in the supplementary table. We feel that the revisions and inclusion of another table [table 3 for regression results] now adds more coherence to data presented in the revised table 2.

Comment 5: A minor comment is that there is no key provided for the superscript letters in Table 2. Presumably these refer to the levels of significance?

**Authors' Response**

Action: Apologies for the typo in the original submission. This error has now been fixed in the revised manuscript.

Comment 6: Out of interest, is there any other SF-36 data from Australian carers? It would be of interest to know if outcomes are different between rural and urban carers.

**Authors' Response**

We agree with the comment. However the information is not available.

Action:

As mentioned above, despite an exhaustive re-checking of all relevant scientific databases as well as grey literature, we have been unable to locate any study that explicitly provides data on SF-36 across urban and rural carers in Australia.

Comment 7: Finally, it would be helpful if the authors could comment on potential self-selection bias in the limitations.

**Authors' Response**

Agree with the reviewer's comment.

Action:

Section on Limitations revised. See p. 18 of the revised manuscript.

**VERSION 2 – REVIEW**

<b>REVIEWER</b>	Sally Chan University of Newcastle, Australia
<b>REVIEW RETURNED</b>	19-Jun-2016

<b>GENERAL COMMENTS</b>	The authors have revised the manuscript according to reviewers' comment. There are still few comments that need to be addressed: 1. Ethical issue - Did those with high scores in CES-D get referral or support? 2. The duration of caregiving and care recipients' stage of illness would have an impact on carers' health status. The authors need to
-------------------------	---

	address this in the discussion and limitation. 3. Please add implications of this study to future research and practice.
--	---

<b>REVIEWER</b>	Peter Watson Medical Research Council UK
<b>REVIEW RETURNED</b>	29-Jun-2016

<b>GENERAL COMMENTS</b>	<p>Self-reported physical and mental health of Australian carers: A cross-sectional study bmjopen-2016-011417.R1</p> <p>The methods are correct although I would like more details on how missing values were treated especially since the reference for how this was done (SF-36 Manual and Interpretation Guide) seems obscure. It is becoming more accepted practice to perform multiple imputation which assumes responses are missing at random and can use predictors, or a constant term, to estimate missing responses. This can be done in SPSS. I also found Table 2 uninformative and confusing.</p> <p>Page 6, line 47. I wasn't clear how imputation was performed on what scores and would have welcomed more detail here. I assume this is single imputation rather than multiple imputation ie filling in missing values using a single (imputation) estimate and then analysing the filled-in data set rather than using a range of estimates and assessing the variability of regression estimates across the (multiple) imputations. SPSS can perform multiple imputation using regression which accounts for variations in estimates of missing variables.</p> <p>Page 8 One could point out (I think!) that the Mahalanobis distance is checking for outliers in the predictor variable values and the Cook's distance outliers in the predicted response to explain why two tests of outliers were used.</p> <p>Page 8, line 2. I wondered why Dunnett's test was used for post-hocs as opposed to the usual more general Bonferroni or Sidak. It is a test, as I understand it, for comparing each group with the control group only and there is no obvious control group here e.g. for performing age band comparisons as in Table 3.</p> <p>Page 8, line 42. Stepwise techniques are used here for the two binary outcomes but on page 13 (lines 46-50) it mentions the multiple regressions for the continuous responses adjust for age, gender, work status and other variables used in the stepwise regression and on page 8 the multiple regressions are not described as being stepwise so I wonder if the same sort of regression (stepwise) was used for both the multiple regressions and logistic regressions to enable comparability. The usual approach adjusting for variables is to enter all the predictors at once to assess how they relate to the response (such as CES-D) independently of each other. I don't see what you gain using a stepwise approach to regression over entering all at once.</p> <p>Page 8, line 47. What was the response rate? Did all people invited to take part participate in the study?</p>
-------------------------	---

	<p>Page 9, The percentages for the primary and secondary caregiver roles only sum to 88% &lt; 100%. What about the other 12%? Similarly the sum of the percentages for caring relationship is 5% below 100% which seems large for rounding error.</p> <p>Page 12, Table 2. The use of asterisks for denoting p-values would be more striking and is the standard. It is also not clear from an inspection of Table 2 exactly what hypothesis tests the p-values are testing and therefore it is difficult to see how, for example, the authors can conclude about the age associations mentioned on page 10, line 55. I suspect it may possibly be comparing groups but then why do males and females both have a p-value in Table 2 for sex with PCS (top left)?</p> <p>Page 14, Table 3. The ASA are suggesting confidence intervals for regression estimates as a useful accompaniment to p-values to show the precision of estimation rather than quoting point estimates so these could be place in Table 3 instead of, or in addition to, the point estimates. Giving confidence intervals also gives an indication of overlap of effects between groups such as the differences across age bands (given in Table 3). Table 3 does not show if there are differences between all pairs of age bands which could be tested other than compared to the reference group. Is there an overall test of group differences for the bandings given in Table 3 e.g. F ratio or chi-square measure?</p>
--	---

### VERSION 2 – AUTHOR RESPONSE

REVIEWER 1 (Sally Chan)

Comments

1. Ethical issue - Did those with high scores in CES-D get referral or support?

Authors' Response:

The survey was largely "online and anonymous" and therefore there was no way we could organise referrals. The Survey Information Sheet which had to be read and ticked prior to starting the online survey clearly indicated that if participants felt anxious or depressed either as a result of the survey or thinking through their caring-giving responsibilities, they should contact the toll-free Lifeline Telephone Support Service, and the relevant telephone number was provided. Lifeline counsellors do suggest contacting their preferred doctor or a psychologist if symptoms have been ongoing.

Text has been added to clarify this situation in the Methodology Section, under the sub-heading of Study Design, setting and participants on page 6 (reproduced below for ease of locating it):

"The majority of the completed surveys (nearly 90%) were done online. As most of the survey responses were on-line and recruitment was undertaken through a variety of community-based strategies, an overall response rate could not be determined. Participation was both voluntary and anonymous. This also meant that there was no capacity to refer individuals potentially at risk for support, however, the information sheet outlining details of the study provided to all potential participants, preceded the beginning of the survey questionnaire, and included specific details to facilitate access to free counselling support if required through Lifeline Telephone Support Service, and the relevant toll-free telephone number was provided. Lifeline counsellors do suggest callers to contact their preferred doctor or a psychologist if symptoms have been ongoing. This strategy was approved by the Human Research Ethics Committee as a viable option to maintain anonymity of participants whilst also providing information on how to access counselling support for those experiencing distress associated with their caregiving role."

2. The duration of caregiving and care recipients' stage of illness would have an impact on carers' health status. The authors need to address this in the discussion and limitation.

Authors' Response:

ACTION: The authors agree with this recommendation, and additional information have been added to the discussion and limitations sections.

Text amended and additional details added to the paper

(see page 17, under Discussion – highlighted section in the manuscript, reproduced below for ease of locating it):

"It is acknowledged that, as an initial study, not all key demographic data was able to be captured. In particular, it is recommended that future research specifically examine two additional factors, the length of time the individual had been providing care, and whether the care recipient was receiving palliative or end-of-life care, as it is acknowledged that these issues may have a further impact upon carers' health."

(see page 19, under Limitations – highlighted section in the manuscript, reproduced below for ease of locating it):

"As noted in the discussion, there were two factors that were not captured in the data, the length of time the individual had been providing care and the care recipients' stage of illness. It is acknowledged that these issues may have consequences for carers' health, but it is not believed that the failure to gain these data invalidates the reported outcomes. Nonetheless, it is recommended that future research consider stratifying the results by these variables to measure if there is any impact."

3. Please add implications of this study to future research and practice.

Authors' Response:

ACTION: The authors agree with this recommendation, and additional information have been added to the discussion and limitations sections.

Text amended and additional details added to the paper

(see page 20, under Conclusion – highlighted section in the manuscript, reproduced below for ease of locating it):

"The need for access to specifically targeted psychological interventions for individuals at risk is clear. Recommendations for future research to build upon these findings include examination of the previously identified factors of length of time the individual had been providing care and the care recipients' stage of illness, as well as undertaking a comprehensive study that comparatively examines carers residing in a variety of rural settings against metropolitan-based peers in order to recognise difference and disparity in health status."

REVIEWER 3 (Peter Watson) - STATISTICAL ADVISOR

The methods are correct although I would like more details on how missing values were treated especially since the reference for how this was done (SF-36 Manual and Interpretation Guide) seems obscure. It is becoming more accepted practice to perform multiple imputation which assumes responses are missing at random and can use predictors, or a constant term, to estimate missing responses. This can be done in SPSS.

Authors' Response:

ACTION: The authors agree that additional details could be useful for the reader. We had a reputable external data management and analyses agency undertake some of the data management and these details have now been included.

Text amended and additional details added to the paper

(see page 7, under Method (data measurement section) – highlighted section in the manuscript, reproduced below for ease of locating it):

"The SF-36 imputation was undertaken by CR&DITSS (Clinical Research Design IT and Statistical Support), which is the statistical consulting arm of the Centre for Clinical Epidemiology and Biostatistics (CCEB) at the Hunter Medical Research Institute affiliated with University of Newcastle. CCEB have an established algorithm for missing values, which has been used by numerous research projects across Australia. Imputation of missing values and computation of the domain and component scores were performed according to the procedures outlined in the SF-36 Manual and Interpretation Guide.<sup>35</sup> The SF-36 domain generation was done using the SAS software. The SAS codes contained in the document [www.sascommunity.org/sugi/SUGI94/Sugi-94-168%20Newvine.pdf](http://www.sascommunity.org/sugi/SUGI94/Sugi-94-168%20Newvine.pdf) was used to create the 9 domains (eight original and the summary component scores for PCS & MCS) using the SF36 variables. For missing data imputation, it was assumed missing data were completely at random and used person-mean imputation for subjects with missing values. Missing values on an item were replaced with the mean of all the individual's completed items. Person-mean imputation has been shown to be reliable when the numbers of respondents with missing items are 20% or less.<sup>36</sup> The imputed data variables were then exported and reinserted in the SPSS data file before undertaking analyses for the present paper."

I also found Table 2 uninformative and confusing.

Authors' Response:

ACTION:

Table 2: This table was changed in response to Reviewer 2's recommendation from the first review. These changes were done precisely as requested, however, the authors' actually agree that this table is now more confusing, and, as Reviewer 2 refused the invitation to review this version, we have chosen to return to the original submitted format in this regard. We have now reverted back to the original table 2 with some minor edits to make it easier to read. Please see page 13 of the revised manuscript.

Page 6, line 47. I wasn't clear how imputation was performed on what scores and would have welcomed more detail here. I assume this is single imputation rather than multiple imputation ie filling in missing values using a single (imputation) estimate and then analysing the filled-in data set rather than using a range of estimates and assessing the variability of regression estimates across the (multiple) imputations. SPSS can perform multiple imputation using regression which accounts for variations in estimates of missing variables.

Authors' Response:

Please see the detailed response above to data imputation.

Page 8 One could point out (I think!) that the Mahalanobis distance is checking for outliers in the predictor variable values and the Cook's distance outliers in the predicted response to explain why two tests of outliers were used.

**Authors' Response:**

The inclusion of two tests (Mahalanobis and Cook's distance) were simply from a perspective of additional clarity for the reader regarding violations. No additional text has been added at this point, however, it would be easily included if considered desirable.

Page 8, line 2. I wondered why Dunnett's test was used for post-hocs as opposed to the usual more general Bonferroni or Sidak. It is a test, as I understand it, for comparing each group with the control group only and there is no obvious control group here e.g. for performing age band comparisons as in Table 3.

**Authors' Response:****Comment**

Dunnett's C was recommended for use with SF-36 data as it provides very tight Type 1 error control and performs well when the group sizes are different, when population variances are different or when data is not normally distributed (e.g. Field, 2009). Reference: Field, A. (2009). Discovering statistics using SPSS. Sage: London.

**ACTION:** Text amended and additional details added to the paper

(see page 8, under Statistical methods – highlighted section in the manuscript, reproduced below for ease of locating it):

"ANOVA with posthoc comparison via Dunnett's C, with  $p \leq 0.05$  as the critical value, were used to test for differences between mean SF-36 scores of participants. Dunnett's C was recommended for use with SF-36 data as it provides very tight Type 1 error control and performs well when the group sizes are different, when population variances are different or when data is not normally distributed."

Page 8, line 42. Stepwise techniques are used here for the two binary outcomes but on page 13 (lines 46-50) it mentions the multiple regressions for the continuous responses adjust for age, gender, work status and other variables used in the stepwise regression and on page 8 the multiple regressions are not described as being stepwise so I wonder if the same sort of regression (stepwise) was used for both the multiple regressions and logistic regressions to enable comparability. The usual approach adjusting for variables is to enter all the predictors at once to assess how they relate to the response (such as CES-D) independently of each other. I don't see what you gain using a stepwise approach to regression over entering all at once.

**Authors' Response:**

**Comment:** We apologize for confusing the reviewer by the sentence structure.

We reviewed all the analyses output files and in fact re-did some of the regression analysis to ensure that there was no error.

**ACTION**

Text has been amended to clarify that the term stepwise was incorrectly used -- when what was meant was "sequential analysis". We have corrected this error. See page 9, last two paragraphs.

Page 8, line 47. What was the response rate? Did all people invited to take part participate in the study?

**Authors' Response:**

Text amended and additional details added to the paper

(see page 6 – highlighted section in the manuscript, reproduced below for ease of locating it):

"The majority of the surveys (nearly 90%) across the state were completed online. As most of the survey responses were on-line and recruitment was undertaken through a variety of community-based strategies, an overall response rate could not be determined."

Page 9, The percentages for the primary and secondary caregiver roles only sum to 88% < 100%. What about the other 12%? Similarly the sum of the percentages for caring relationship is 5% below 100% which seems large for rounding error.

Authors' Response:

ACTION

Text amended and additional details added to clarify the issue (see page 9 – first paragraph of Result section - highlighted section in the manuscript, reproduced below for ease of locating it):

"Raw percentages rather than cumulative percentages are reported; as there are missing values the final totals do not necessarily tally to 100%."

Page 12, Table 2. The use of asterisks for denoting p-values would be more striking and is the standard. It is also not clear from an inspection of Table 2 exactly what hypothesis tests the p-values are testing and therefore it is difficult to see how, for example, the authors can conclude about the age associations mentioned on page 10, line 55. I suspect it may possibly be comparing groups but then why do males and females both have a p-value in Table 2 for sex with PCS (top left)?

Authors' Response:

Action: The authors agree that asterisks allow more clarity and Table 2 has been changed accordingly. As acknowledged earlier, changes in the revised paper in response to comments by Review 2 resulted in Table 2 becoming both unclear and confusing. We have reverted to our original (and preferred) version of Table 2, with some minor edits as noted elsewhere.

Page 14, Table 3. The ASA are suggesting confidence intervals for regression estimates as a useful accompaniment to p-values to show the precision of estimation rather than quoting point estimates so these could be placed in Table 3 instead of, or in addition to, the point estimates. Giving confidence intervals also gives an indication of overlap of effects between groups such as the differences across age bands (given in Table 3). Table 3 does not show if there are differences between all pairs of age bands which could be tested other than compared to the reference group. Is there an overall test of group differences for the bandings given in Table 3 e.g. F ratio or chi-square measure?

Authors' Response:

ACTION

Table 3 has been amended to include 95% Confidence Interval data in addition to p value. See page 16 for amended Table 3.