

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

TITLE (PROVISIONAL)	A cross-sectional study of resilience, burnout and coping mechanisms in UK doctors.
AUTHORS	McKinley, Nicola; McCain, R; Convie, Liam; Clarke, Mike; Dempster, Martin; Campbell, William; Kirk, Stephen

VERSION 1 – REVIEW

REVIEWER	Linda Perkins-Porras St George's, University of London England UK
REVIEW RETURNED	10-Jun-2019

GENERAL COMMENTS	<p>Well written interesting paper on an important and emerging topic area. Introductions sets context concisely but in good detail. Aims clearly set out. Methods described well and study appears to have been conducted to a good standard. Outcomes and results presented clearly. Interesting and relevant discussion. This study adds a valuable contribution to informing the current debate of burnout and stress management among medical staff. Findings support other studies investigating these issues.</p> <p>Tables need to be formatted correctly, should have individual descriptive titles and would benefit from providing the range of scores for each measure to remind reader.</p>
-------------------------	---

REVIEWER	Erin Dehon, PhD University of Mississippi Medical Center
REVIEW RETURNED	20-Jun-2019

GENERAL COMMENTS	<p>Overall, I believe this is an important topic/area of research and the authors were able to recruit a relatively large sample of physicians (though the gender breakdown is disappointing). The manuscript could be strengthened by providing more detail about the results and specific scores for the measures. Right now the results are a bit disorganized and confusing. There is a lot of mention of mean scores for measures without any description of what those scores indicate.</p> <p>Abstract:</p> <ul style="list-style-type: none">• Mean scores without a description of what they imply is not very informative <p>Introduction.</p> <ul style="list-style-type: none">• The goal/objective of the study is not clearly stated• Is this study simply an extension of the pilot study—larger sample size <p>Methods</p>
-------------------------	---

	<ul style="list-style-type: none"> • I would like more information about the recruitment process. Please write this section in way so others could repeat exactly what you did. For instance, the authors currently state “Royal Colleges and other medical organizations...” Who are “other medical organizations”? How many potential participants were invited to participate? This will allow the authors to calculate a response rate. Even if you cannot calculate response rate it would be valuable to have an idea about how many physicians using each method (email, social media) were invited • Did you ensure that the survey was not completed more than once by the same physician since presumably the same physician could see the survey on an email invitation and on social media. • In the second paragraph of the methods, rather than mentioning using 3 “psychological tests” state something like 3 empirically validated instruments that assess resilience, professional quality of life, and coping mechanisms. Then I would go into detail about each measure, its subscales, responses scale and psychometric properties. A description of each scale may take a paragraph each including a description of the score ranges and interpretation of scores. For example the authors discuss low, medium, and high burnout and secondary stress in the results—low medium and high ranges should be clearly identified in the methods section. • I recommend eliminating the term “psychological tests” throughout the paper • Remove paragraph on free text response options that were not reported in this paper <p>Results.</p> <ul style="list-style-type: none"> • Why did so few males participate? How does the breakdown by gender compare to the population of physicians in the area? • Mean scores without a description of what they imply is not very informative <p>Discussion.</p> <ul style="list-style-type: none"> • Findings that burnout is highest among Emergency Medicine physicians is consistent with findings of physicians in the U.S.—authors should reference this. • Add lack of male participation to limitations
--	---

REVIEWER	David Culliford University of Southampton
REVIEW RETURNED	25-Jul-2019

GENERAL COMMENTS	<p>I congratulate the authors on an interesting and important manuscript entitled ‘National study of resilience, burnout and coping mechanisms in UK doctors’. What follows is a statistical review of the manuscript.</p> <p>Methods</p> <p>Page 7, line 37: Please state the name of the survey software which was used.</p> <p>Page 8, lines 34-41: You make no mention here of any non-parametric tests or methods being used, so I assume that you assessed the data for normality before using the methods stated (i.e. t test, ANOVA, linear regression). Please add a covering sentence to say something like “the assumption of normality was assessed by visual means” or similar.</p> <p>Results</p> <p>Page 9, lines 14: If you are using the phrase “... no significant difference ...” then you ought to add the test statistic and p-value (for the difference in resilience between men and women) to support this statement.</p>
-------------------------	---

	<p>Page 9, lines 19-22: Here, and later on in the results section, you have made a statement about an ANOVA result. For this test, a statistically significant result (from an F test) does not necessarily imply that any one category (or subgroup of categories) is significantly different from any other category subgroup. You are probably fine here since you are avoiding the word “significantly”, but just be mindful about the wording and any inferences the reader may draw. If you really wanted to make a statement of significance difference between any two groups of category levels, then a post-hoc test would be required, along with the requisite adjustment for multiplicity.</p> <p>Page 11, lines 9-15: Please provide test statistics and p-values for each of the results reported – presumably these were t tests against a (fixed) population mean.</p> <p>Page 11, lines 9-15: Ditto, see previous comment on ANOVA (Page 9, lines 19-22). Just because you get a significant F test result from an ANOVA, the result itself does not necessarily imply that it is solely the most outlying category (Emergency Medicine) that is different from all other groups. Yes, in this case it almost certainly is Emergency Medicine that is driving the F test’s significant p-value. However, the wording is probably OK as it stands.</p> <p>Page 14, line 22: Please consider removing the p-value. The statement is fine without it and the fact that substance abuse as a coping strategy was infrequently reported makes the stated p-value seem odd within that sentence.</p> <p>Discussion</p> <p>General comment: I like the way that you have refrained from making any strong inferential statements about the results, and that you are clear about this in the study limitations paragraph</p> <p>Page 15, line 31: I think that these two reference numbers are not quite right (should be numbers 22/23, and not 23/24?) – I checked the articles referenced and the CD-RISC summary statistics seem to reflect this.</p> <p>Tables</p> <p>For tables 1/2/3, please consider rewording the title for the CD-RISC column as: Conor Davidson Resilience Scale (CD-RISC) (Mean, SD)</p> <p>For all tables, please consider: Please review whether you need the number of decimal places (d.p.) used for the various scores reported in each of the tables. One d.p. (two at most) should be sufficient for each of the scales used. Also please ensure that the number of decimal places is consistent within each table.</p> <p>For Table 6: It would be nice to add a footnote to this table explaining that the regression used was backwards stepwise elimination and then list all the variables entered into the regression model at the first step. This is just for transparency. Also please update the STROBE statement to reflect this.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Linda Perkins-Porras

Institution and Country: St George's, University of London

England
UK

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below

Well written interesting paper on an important and emerging topic area. Introduction sets context concisely but in good detail. Aims clearly set out. Methods described well and study appears to have been conducted to a good standard. Outcomes and results presented clearly. Interesting and relevant discussion. This study adds a valuable contribution to informing the current debate of burnout and stress management among medical staff. Findings support other studies investigating these issues.

Tables need to be formatted correctly, should have individual descriptive titles and would benefit from providing the range of scores for each measure to remind reader.

Individual descriptive titles have been added to each table.

The range of scores has been included at the bottom of each table to remind the reader.

Reviewer: 2

Reviewer Name: Erin Dehon, PhD

Institution and Country: University of Mississippi Medical Center

Please state any competing interests or state 'None declared': none declared

Please leave your comments for the authors below

Overall, I believe this is an important topic/area of research and the authors were able to recruit a relatively large sample of physicians (though the gender breakdown is disappointing). The manuscript could be strengthened by providing more detail about the results and specific scores for the measures. Right now the results are a bit disorganized and confusing. There is a lot of mention of mean scores for measures without any description of what those scores indicate.

Abstract:

- Mean scores without a description of what they imply is not very informative
Introduction. This has been amended with population norm scores for each scale included.

- The goal/objective of the study is not clearly stated.
The final paragraph of the introduction now states the aims clearly.

It aimed to investigate resilience, stress, burnout and coping in doctors across the UK. It also aimed to determine the impact of demographic variables, such as sex, grade and speciality on these factors. Finally, it aimed to assess the factors most strongly related to burnout in UK doctors. As far as we are aware, it is the first study to report on these factors in the UK's medical workforce.

- Is this study simply an extension of the pilot study—larger sample size The pilot study was a single NHS trust in Northern Ireland as stated in methods. This study includes all UK doctors.

Methods

- I would like more information about the recruitment process. Please write this section in way so others could repeat exactly what you did. For instance, the authors currently state "Royal Colleges and other medical organizations..." Who are "other medical organizations"? All Royal Colleges and medical organisations that promoted the survey have been listed in appendix 1. Appendix 1 also details how each organisation choose to advertise the survey to potential participants.

How many potential participants were invited to participate? This will allow the authors to calculate a response rate. Even if you cannot calculate response rate it would be valuable to have an idea about how many physicians using each method (email, social media) were invited.

It is impossible to know how many doctors saw the invitation to the survey as there is no information on how many doctors use social media accounts such as Twitter.

- Did you ensure that the survey was not completed more than once by the same physician since presumably the same physician could see the survey on an email invitation and on social media. Participants were asked to complete the online questionnaire once and only if they were a medical doctor currently working within the UK.

- In the second paragraph of the methods, rather than mentioning using 3 “psychological tests” state something like 3 empirically validated instruments that assess resilience, professional quality of life, and coping mechanisms.

This has been amended.

Then I would go into detail about each measure, its subscales, responses scale and psychometric properties. A description of each scale may take a paragraph each including a description of the score ranges and interpretation of scores. For example the authors discuss low, medium, and high burnout and secondary stress in the results—low medium and high ranges should be clearly identified in the methods section. This has been amended with a separate paragraph detailing each measure.

Additionally, each component can be transformed to a t-score and divided into low (≤ 43), moderate (44–56) or high (≥ 57) to facilitate interpretation of scores.

- I recommend eliminating the term “psychological tests” throughout the paper.

This has been amended.

- Remove paragraph on free text response options that were not reported in this paper

This has been removed.

Results.

- Why did so few males participate? How does the breakdown by gender compare to the population of physicians in the area?

More male than female doctors are registered with the General Medical Council in the UK. However, it is typical that more females than males participate in online surveys(referenced).

- Mean scores without a description of what they imply is not very informative

Population mean scores are now highlighted throughout the manuscript.

Discussion.

- Findings that burnout is highest among Emergency Medicine physicians is consistent with findings of physicians in the U.S.—authors should reference this. This has been amended.

The following studies have now been referenced:

1. Shanafelt TD, Boone S, Tan L, et al. Burnout and satisfaction with work-life balance among US physicians relative to the general US population. *Arch Intern Med* 2012; 172:1377–85.

2. Peckham C. Medscape National Physician Burnout & Depression Report 2019. Medscape. 2019. [Accessed 23 August 2019]. Available at: <https://www.medscape.com/slideshow/2019-lifestyle-burnout-depression-6011056#2>.

- Add lack of male participation to limitations This has been amended.

Reviewer: 3

Reviewer Name: David CULLiford

Institution and Country: University of Southampton

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

I congratulate the authors on an interesting and important manuscript entitled 'National study of resilience, burnout and coping mechanisms in UK doctors'. What follows is a statistical review of the manuscript.

Methods

Page 7, line 37: Please state the name of the survey software which was used. This has been amended.

The survey was created using survey software, SurveyMonkey[®] and a hyperlink was generated.

Page 8, lines 34-41: You make no mention here of any non-parametric tests or methods being used, so I assume that you assessed the data for normality before using the methods stated (i.e. t test, ANOVA, linear regression). Please add a covering sentence to say something like "the assumption of normality was assessed by visual means" or similar.

The assumption of normality was assessed by visual means and this has been added to the manuscript.

Results

Page 9, lines 14: If you are using the phrase "... no significant difference ..." then you ought to add the test statistic and p-value (for the difference in resilience between men and women) to support this statement. This has been amended.

Page 9, lines 19-22: Here, and later on in the results section, you have made a statement about an ANOVA result. For this test, a statistically significant result (from an F test) does not necessarily imply that any one category (or subgroup of categories) is significantly different from any other category subgroup. You are probably fine here since you are avoiding the word "significantly", but just be mindful about the wording and any inferences the reader may draw. If you really wanted to make a statement of significance difference between any two groups of category levels, then a post-hoc test would be required, along with the requisite adjustment for multiplicity.

Page 11, lines 9-15: Please provide test statistics and p-values for each of the results reported – presumably these were t tests against a (fixed) population mean. This has now been included.

Page 11, lines 9-15: Ditto, see previous comment on ANOVA (Page 9, lines 19-22). Just because you get a significant F test result from an ANOVA, the result itself does not necessarily imply that it is solely the most outlying category (Emergency Medicine) that is different from all other groups. Yes, in this case it almost certainly is Emergency Medicine that is driving the F test's significant p-value. However, the wording is probably OK as it stands.

Page 14, line 22: Please consider removing the p-value. The statement is fine without it and the fact that substance abuse as a coping strategy was infrequently reported makes the stated p-value seem odd within that sentence. This has been removed.

Discussion

General comment: I like the way that you have refrained from making any strong inferential statements about the results, and that you are clear about this in the study limitations paragraph

Page 15, line 31: I think that these two reference numbers are not quite right (should be numbers 22/23, and not 23/24?) – I checked the articles referenced and the CD-RISC summary statistics seem to reflect this. This has been amended.

Tables

For tables 1/2/3, please consider rewording the title for the CD-RISC column as:

Conor Davidson Resilience Scale (CD-RISC)

(Mean, SD) This has been amended.

For all tables, please consider:

Please review whether you need the number of decimal places (d.p.) used for the various scores reported in each of the tables. One d.p. (two at most) should be sufficient for each of the scales used. Also please ensure that the number of decimal places is consistent within each table. This has been amended. All tables now one d.p. (excluding table 6.)

For Table 6:

It would be nice to add a footnote to this table explaining that the regression used was backwards stepwise elimination and then list all the variables entered into the regression model at the first step. This is just for transparency. Also please update the STROBE statement to reflect this. This has been amended, as has the STROBE statement.

VERSION 2 – REVIEW

REVIEWER	Erin Dehon PhD University of Mississippi Medical Center United States
REVIEW RETURNED	26-Sep-2019

GENERAL COMMENTS	The authors have done an excellent job addressing all of the reviewers concerns. I don't have any further suggestions.
-------------------------	--

REVIEWER	David Culliford University of Southampton, U.K.
REVIEW RETURNED	24-Sep-2019

GENERAL COMMENTS	You have done a thorough job of addressing the reviewers' comments and I congratulate you on a fine manuscript. I am recommending acceptance for publication.
-------------------------	---