PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Need for recovery amongst Emergency Physicians in the United Kingdom and Ireland: a cross-sectional survey
AUTHORS	Cottey, Laura; Roberts, Tom; Graham, Blair; Horner, Daniel; Stevens, Kara; Enki, Doyo; Lyttle, Mark; Latour, Jos

VERSION 1 – REVIEW

REVIEWER	Kristine Olson Yale School of Medicine, USA
REVIEW RETURNED	08-Jul-2020

GENERAL COMMENTS	Dear authors, Thank you for the great privilege to read and review your work. The paper was well written. I had confidence in the work. I think it adds new and important information to the existing literature. I offer the following suggestions for your consideration. Abstract: As a standalone document, the abstract is not clear. This is mostly because “need for recovery” is not clearly defined. In my opinion, the conclusions provided on page 16 Lines 337 to 341 provide more clear results for the abstract. Also, page 18 Lines 371 to 374 provide more clear conclusions. Introduction: Page 6 lines 125 to 126: It is as though you hypothesize that the unscheduled care setting and high intensity work as seen in emergency medicine makes it difficult to recovery between shifts. Another way of looking at this might be that the lack of control over volume and intensity creates the need for recovery. I think these concepts tie in with your conclusions that lack of control over time to recover is correlated with high “need for recovery” scores. I suggest introducing the need for recovery scale, and the definition of need for recovery, early on in the paper. The narrative of the work is dependent on having a conceptual framework for NFR. Methods: The methodology is sound and gives me confidence in the work by these authors. It is very well described, easy to follow, and easy to repeat. The sample and sample size are outstanding. The inclusion exclusion criteria is good with the following considerations: the

	sample includes emergency physicians, non-emergency physicians, and trainees. I question how similar or different these groups are, and how generalizable the work is to emergency medicine versus exposure to emergency medicine. The sample size is 29% ages 26 to 30 years old. I question whether these are trainees, and whether they are trainees specific to emergency medicine. Trainees from all specialties are known to have a higher proportion of burnout, and likely have a greater need for recovery. You might consider excluding trainees or those who are not specifically emergency medicine physicians. I understand this would reduce your sample size, but you may get a clear understanding of work life in emergency medicine. If you keep non-emergency medicine physicians and trainees in this group, I would describe the nature of their work and consider the influence on the interpretation in the discussion. Also, it is not clear to me how the models were adjusted. Were they adjusted for age and gender, and only those factors that were independently significant? Where they controlled for whether of not you were EM, GP/consultant, trainee? I would clarify. The methodology and statistical analysis are so great and well described it almost becomes the predominate feature of the paper. I would use this paper to teach STATA. In general, I would review the paper for emphasizing the narrative importance of the work. Need for recovery scale – I would describe the scale, as it is necessary to understanding the paper. I am glad you provided the survey. I would refer to it. Also, two of the 11 items are oriented in a different direction. Have those been reversed in scoring? Also, how are the 11 yes-no items scored such that you can have an median score of 70? In the six-week timeframe for the data collection, are the invitees prompted with any specific periodicity (ex. weekly)? I would read the manuscript from the perspective of someone less familiar with the work and strengthen the narrative by providing definitions and descriptions where needed for those less familiar with the concept of “need for recovery” and the nature of the work of those included in the sample. Results: Table 2: The fact that you have lower “need for recovery scores” the more months that you were in the emergency department makes me question whether this was because they had more experience in emergency medicine which makes you more comfortable and less burdened out by the work, or, if the sheer fact that the trainees account for those working less than a year in the emergency department. These findings may or may not be different if you exclude trainees from the sample. Where is, how much time you spend in the emergency department based on your contract does not follow the same pattern. However, greater exposure based on number of weekends worked and number of consecutive days worked does seem related to a greater need for recovery. It was not surprising that there is less need for recovery with increasing age, nor male gender. I was intrigued by the fact the

	consultants/GPs had lower need for recovery scores then the emergency medicine physicians. Again, I'm curious how old these roles are similar or different. Again, I wonder if including only emergency medicine physicians would give you a clearer understanding of emergency medicine work. If you include the consultants and the trainees, I think it's important to explain why you feel justified to do so. I think your findings are very important, especially finding that greater exposure to work with number of consecutive shifts, proportion of weekends worked, and more work after work are all associated with the need for recovery, or that the ability to recover is impaired. I think it's an important contribution to find that control over time to recover (study or leave) may be protective, reduces the need for recovery. I think this paper would be frequently cited for these findings. I would be clear about the models and adjustments. For your international audience, I would better describe the meaning of "clinical grade" and non-emergency medicine doctor "consultant'. In the United States we generally do not have GPS or consultants staffing the emergency room. For your international audience, you might want to be clear about what it means "out of work hours", some might say "after hours work" or "work after work" or "work outside of work hours". I would have liked to the need for recovery scores for dayshifts versus night shifts. I see it as part of your survey. I would've been interested to see how these findings "need for recovery" compare with your question "I am currently suffering burn out from work" and "I feel it higher risk of burn out from my job in the near future". Part of your motivation in this work was to determine if the "need for recovery scale" would likely identify the need for recovery sooner than assessing burnout. Perhaps this will be a future paper. However, at this point I'm not convinced that there is a need for the new scale such as "need for recovery" if those other metrics give you the same results, as this additional scale takes up valuable real estate in precious wellness surveys. Yet, I do think this paper is valuable such that the scientific community can consider this notion. This simple 11 question "need for recovery scale" is interesting. I also see that you have assessed the amount of staff support per respondent – I look forward to seeing your future paper on whether or not the amount of adjunct support decreases the need for recovery. I agree with you, it was strange that you found those with "significant caring responsibilities outside of work" have a lower "need for recovery" score. Overall, I thought the paper was very well done. It was a well written. The methodology and statistics or well done and well described. (Please make the models and adjustments clearer.) I think it adds a valuable new metric and findings to the existing literature. I especially find the results are valuable in a practical sense for changing policies and practices. The two things I would consider most seriously is who is included in the sample for the hypothesis you wish to test, I would strengthening the narrative with definitions and descriptions to tell the story of the science.

REVIEWER	Erin Dehon University of Mississippi Medical Center U.S.
REVIEW RETURNED	09-Jul-2020

GENERAL COMMENTS	Interesting study. Methods and results are well described. Major strengths--sample size and response rate Major weakness--lack of other well being related measures. It would be helpful to see how NSF score relate to outcomes such as burnout and fatigue among EPs. We can infer this from previous research using participants in other occupations but as this study shows EPs are responding quite differently (more extreme scores) compared to individuals in other occupations. I do wish the authors would make a stronger argument for the use of this scale among EPs. Why should this measure be use compared to existing measures (e.g., burnout, fatigue)? The introduction states that burnout measures lack the ability to define specific contributory factors or highlight opportunities for intervention. I don't understand this argument or perhaps need more clarification. I do not see how the NSF is any better than existing burnout instruments in terms of the information it provides. Previous studies using burnout measures have similarly (as in this study) compared burnout to modifiable factors such as access to annual leave, hours spent charting after work, etc. This manuscript could be strengthened by highlighting the strengths of the NFR compared to other similar measures, including but not limited to fatigue measures. Also how is the NFR different from other measures. There is a pretty large difference in scores between EPs and other occupations. This is briefly noted in the discussion. I would recommend highlighting this finding a bit more. Perhaps reiterating in the discussion the EP median score of 70 compared to that of other samples (mean = 36-44). This finding warrants a bit more discussion.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Thank you for your in-depth review and your comments on the manuscript. In response to your suggestions and questions we have addressed the following issues in the revised manuscript:

1. As a standalone document, the abstract is not clear

This has been revised in line with suggestions, page 3 lines 80-81, 84-89, 97-102.

2. I suggest introducing the need for recovery scale, and the definition of need for recovery early on in the paper.

The introduction has been revised in line with suggestions. In addition, we have further revised the introduction to emphasise the difference of the NFR scale to other measures of burnout/fatigue and

strengthen the overall narrative of the work, in line with your other comments and those of Reviewer 2.

3. If you keep non-emergency physicians and trainees….I would describe the nature of their work and consider the influence on the interpretation in the discussion

We have provided a clearer definition of the different physician roles within the methods, page 8 line 212-219, to assist with understanding of the inclusion in the study. The discussion has also been revised to reflect this suggestion, page 19, line 437-440 and additional online supplementary material to clarify the training structure in the UK matched to that of North America.

4 . It is not clear to me how the models were adjusted?...I would clarify

We have revised the statistical analysis section and added a footnote to Table 4 to emphasise model adjustment and the process used. Each coefficient estimate is adjusted for all other covariates which were statistically significant in the model.

5. Need for recovery scale – I would describe the scale… and strengthen the narrative by providing definitions and descriptions

We have revised the introduction in line with earlier suggestions (point 2) and explaned the explanation of the Need for Recovery scale and its benefits within the methods (page 9, lines 227-233) and discussion (page 21, lines 469-473).

6. Also, how are the 11 yes-no items scored such that you can have an median score of 70?

Scores were summated to give a score out of 100, however some scores had to be imputed as described in page 11, lines 281-283 giving a median NFR of 70.

7. were invitees prompted with any specific periodicity?

Weekly reminders were sent out by email and local reminders encouraged through site PIs. We have added a sentence to the methods (page 11, lines 267-268) to reflect this.

8. If you include the consultants and trainees, I think It's important to explain why you feel justified to do so

Please see response to point 3.

9. I would be clear about the models and adjustments

Please see response to point 4.

10. For your international audience… you might want to be clear about what it means 'out of work hours'.

Thank you for highlighting this. We have now clarified what we mean by a 'out of work hours' by providing a description within the methods section (page 11, lines 251-253).

11. Part of your motivation in this work was to determine if the 'need for recovery scale' would likely identify the need for recovery sooner than assessing burnout.

As you correctly suggest, we have another paper in draft stage looking at the relationship between higher need for recovery scores and subjective perception of burnout. This is mentioned briefly in the discussion section. However, in this article we sought to assess the baseline NFR and the influence of potentially modifiable factors, with the assumption that NFR can be a precursor to formal definitions of burnout

12. Overall, I thought this paper was very well done

Thank you

Reviewer 2

Thank you for your in-depth review and your comments on the manuscript. In response to your suggestions and questions we have addressed the following issues in the revised manuscript:

1. It would be helpful to see how NFR score relates to outcomes such as burnout and fatigue among EPs

We plan to report this in a separate paper, as mentioned in the response to point 10 from reviewer 1. Given your interest, we have now added this issue to the discussion and highlight how further work resulting from this project will address this question in the near future.

2. I do wish the authors would make a stronger argument for the use of this scale among EPs

We have extensively revised the introduction in line with this comment (page 6, lines 161-169) and points made by reviewer 1 above (2 and 5), which we hope make a clearer case for use of NFR. We have also included further justification within the methods relating to our PPI consultation (page 10, lines 244-248) and the discussion (page 21, lines 469-475)

3. There is a pretty large difference in scores between EPs and other occupations.

We were also interested to see this. We have already highlighted this in the discussion but have now expanded this slightly to include potential reasons for discrepancy.

VERSION 2 – REVIEW

REVIEWER	Erin Dehon University of Mississippi Medical Center US
REVIEW RETURNED	11-Aug-2020

GENERAL COMMENTS	Authors contributed significant improvements to the introduction Minor comments: How do the 11 yes/no items produce a score from 0-100? In the discussion the authors state "We have identified simple interventions that may reduce NFR" This makes it sound like this particular study focused on identifying interventions to reduce NFR.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1

1. How do the 11 yes/no items produce a score from 0-100?

We have amended the methods section, page 9, line 234 to detail exactly how the score is produced. This is done by taking the total sum of the unfavourable responses, multiplying by 100 and dividing by the number of scale items which is 11.

2. In the discussion the authors state "We have identified simple interventions that may reduce NFR" This makes it sound like this particular study focused on identifying interventions to reduce NFR.

We agree with your observations on this section and to avoid any confusion, we have delete this sentence.

Thank you very much