

Article details: 2020-0068	
Title	<b>Global prevalence of burnout among Postgraduate Medical Trainees: a systematic review and meta-regression</b>
Authors	Leen Naji MD, Brendan Singh MD, Ajay Shah BSc, Faysal Naji MD, Brittany Dennis MD PhD, Owen Kavanagh MD, Laura Banfield MLIS, Akram Alyass PhD, Fahad Razak MD PhD, Zainab Samaan MBChB PhD, Jason Profetto MD, Lehana Thabane PhD, Zahra N. Sohani, MD PhD
<b>Reviewer 1</b>	Leanna Graham
Institution	Toronto General Hospital, Collaborative Academic Practice Director Professional Practice & Policy
General comments and author response	<p>Comments to the Author</p> <p>Reviewer 1 Comment 1: Thank you for the opportunity to review and your attention to this important matter as HCWs around the world focus tirelessly on Covid-19 which is impacting their wellbeing.</p> <p>The results are not necessarily surprising given the focus of recent efforts on culture, psychological safety and the relationship between burnout with errors and quality of care. I note you reference some authors connected with IHI Joy in Work and National Academy of Medicine Collaborative. Would suggest further integration of these concepts into your analysis and conclusions otherwise the simple review of the literature is not as relevant in the current climate and context and will not address your objective to identify risk and protective factors of burnout.</p> <p>Namely:</p> <ul style="list-style-type: none"> <li>-psychological safety to be vulnerable as a clinician, and particularly physician and trainee to discuss burnout</li> <li>-need to change culture at the systematic level which you only briefly describe</li> <li>-deepen your analysis that the approach to address is multifactorial, well documented up to 80% of burnout is system and culture; and only 20% individual</li> <li>-some context of Canadian Universities studies of their residents such as U of T Voice of the Resident.</li> <li>-reflect the relationship between clinician burnout with errors and quality of care and patient experience. The Quadruple Aim is well described by IHI as well as Health Quality Ontario</li> </ul> <p><b>Author Response to Reviewer 1</b> Comment 1: We would like to thank the reviewer for taking the time to thoroughly review and reflect on our manuscript. We completely agree that there are multiple contributing factors to burnout and that systemic issues play a significant role. We have revised parts of our discussion to elaborate on this on pages 8-9, while being mindful of our word limit.</p> <ul style="list-style-type: none"> <li>• We overview the need to change culture at the systemic levels at multiple points in the document. On Page 8: “our findings nonetheless warrant future research to identify cultural and systemic differences that may explain our results, both within and outside the training environment. This hypothesis is supported by other recent work; a cross-sectional study reports that physicians as a group are more resilient than the general population lending less credence to the view that individual factors lead to burnout.(31) Furthermore, a systematic review by Panagioti et al., evaluating strategies to mitigate burnout emphasized the need for organizational level change.(4)”</li> <li>• We highlight some initiatives taken to address systemic limitations of the medical culture on Page 8: “To this end, considerable qualitative work has focused on understanding the pitfalls of medical training, the hidden curriculum, and challenges within the medical culture. Some training programs have sought to overcome the potential toxic culture by incorporating mentorship programs to promote collegiality(27) and create platforms to give voice to PMTs(28). It is likely that answers lie at the intersection of further quantitative research of structural differences between geographically diverse training systems and</li> </ul>

	<p>qualitative work understanding the prevailing toxic culture of medicine and its impact on physicians and patients alike.(29)”</p> <ul style="list-style-type: none"> <li>• Additionally, we added a section titled “Future Direction”, under which we elaborate on the need for systemic change. See page 9. As the reviewer suggests, we also comment on the impact of burnout on the patient safety and productivity in our introduction: “The effects of burnout are widespread, impacting both physician wellness and productivity as well as patient health outcomes.(2,4)”</li> <li>• We also discuss that the problem is multifactorial and draw upon data from other fields to understand what deep rooted differences may exist between geography and work culture overall that may be contributing to burnout. On Page 8: “While there is a paucity of research comparing health care systems between these regions, a study of general workplace trends finds that factors such as more involved unions and longer paid vacations, among other such social policies, contribute to overall improved work-life balance and less burnout.(30)”</li> <li>• The purpose of this systematic review is to quantify burnout over time and identify risk and protective factors. We provide a discussion surrounding our findings with regards to our interpretation of them and in the context of the current literature. We do intend to conduct future studies to develop interventions which may help to mitigate the risk of burnout, for which the resources outlined by the reviewer would provide an incredible wealth of information. However, given our word limit and the current objectives of this study, we are unfortunately unable to provide a more in-depth evaluation of these resources.</li> </ul>
<b>Reviewer 2</b>	Shelly-Anne Li
Institution	
General comments and author response	<p>[22] Reviewer 2 Comment 1: Thank you for the authors' work on an important topic in medical education. Here are several comments for their consideration:</p> <p>Introduction: Previous systematic reviews reported on the prevalence of resident burnout; please include previous reviews and clearly delineate the reasons for performing another review (these other reviews include JAMA's Rotenstein et al., 2018 and PLOS One's Rodriguez et al., 2018); please explain how the aims of your review are different from other reviews, and the relevance of answering your review questions.</p> <p><b>Author Response to Reviewer 2</b> Comment 1: First, we would like to thank the reviewer for taking the time to review our manuscript and providing constructive feedback. While several reviews have aimed to delineate rates of burnout among resident physicians, they have been limited in their search terms thus restricting them to a primarily North American population. Their search terms were restricted to “resident”, excluding many other countries where postgraduate medical trainees are referred to as “junior doctor”, “house officer”, “training physician” etc. For instance, the Rodriguez et al study the reviewer refers to included 26 studies, compared to ours which included 196 studies. Additionally, the Rotenstein et al study was limited to attending physicians, not training physicians. Furthermore, our study identifies using data derived in a robust and systematic manner that interventions aimed at reducing burnout over the past 2 decades have not been successful. The previous reviews have not been able to provide this insight nor have studied variation by geography. We thank the reviewer for encouraging us to highlight the difference between our study and those that came before us; we have revised the manuscript to highlight this in both our introduction, page 3: “Prevalence of burnout among PMTs vary extensively from 3 to 88 percent, but existing attempts at systematic investigation of burnout in this group are limited by methodological flaws including restrictive search terms, lack of evaluation of temporal and associated factors, as well as local, not global, investigation. 2–6”, and in the strengths section of our manuscript: page 9 “Our systematic review employed broad search terms to capture data from 44,128 PMTs across 47 countries. There have certainly been previous systematic reviews on this topic. However, these have been restricted to a largely North American population, included far</p>

fewer studies, or focused on attending physicians, excluding trainees.<sup>1,4–6</sup> The comprehensiveness of our data make our results generalizable and provides a solid platform on which additional data can be added to make more robust conclusions.” We have also added cited the previous systematic reviews in our manuscript for the readers’ reference.

[23] Reviewer 2 Comment 2: Consider removing 'global literature' from page 4, line 39, as this implies the authors have reviewed literature beyond English (which was not the case). There are many reviews that are published in journals that are non-English (e.g., There are 2500 biomedical journals in China; 5% of which are indexed in MEDLINE (Xia et al., 2008), and this was over 10 years ago. Other biomedical journals that do not tend to have their articles translated include journals from Iran, Spain, and India). Also, it is implied that when a systematic review is conducted in North America, that any articles from the main databases we use (MEDLINE, EMBASE, ERIC, PsycINFO, Web of Science) will include all journals we can access in English, and will be providing findings based on such searches (without using the term 'global').

Author Response to Reviewer 2 Comment 2: Thank you for highlighting this important distinction. We have removed the term global from the suggested page and line. As our data comes from 47 countries, we simply state that the data is derived from countries across the world. Additionally, we elaborate on this limitation on page 9 as discussed in our response to [14].

[24] Reviewer 2 Comment 3: Data extraction and quality assessment (Methods): On page 5, line 34, you indicated that there is a lack of validated tools to assess cross-sectional studies (thus prevents you to assess the quality of articles systematically). The CLARITY group from McMaster University provided a risk of bias tool that is specifically for cross-sectional surveys on attitudes and practices. They provided clear guidance on how to assess risk of bias in cross-sectional studies in a systematic way. Why isn't this tool/guide used? {Editorial note: This is an interesting point, but you do not need to use a different RoB tool. Please note that you did not use this tool.}

Author Response to Reviewer 2 Comment 3: Thank you for your feedback. There are many risk of bias tools that have been proposed to assess cross-sectional studies. These tools usually follow a similar framework focusing on known biases (selection, sampling, etc) known in this type of study design. We adapted the Newcastle Ottawa Scale, as has been commonly done in systematic reviews of cross sectional studies, such as in Rotenstein et al.'s systematic review of burnout in attending physicians published in JAMA in 2018.<sup>1</sup> Please see our response to [11] where we address our assessment of risk of bias. The CLARITY group has also developed a similar framework, though it has yet to become widely adapted. Nonetheless, we follow a similar framework and report on individual domains that affect risk of bias as opposed to an overall risk of bias score, as they suggest. Additionally, the tool developed by the CLARITY group is more applicable to assess RoB in studies where the tools are not yet validated, as it includes a section on “pilot testing” of the survey instrument. In our study, the majority of the studies used validated instruments which would not have required pilot testing and these studies would therefore be unnecessarily marked as being at high risk of bias. Contrastingly, our tool includes a component on outcome reporting which is not included in the CLARITY tool but is important for assessing RoB in our study. This is the primary difference between the tools. Therefore, we feel that the modified version of the Newcastle Ottawa Scale, which is the more commonly adapted tool in cross-sectional studies, is more suitable for our review. We have revised the manuscript on page 4 to elaborate on this and have cited the tool used.

[25] Reviewer 2 Comment 4: On page 5, line 36 you indicated, 'However, using the general framework of well-established tools the same reviewers..' Please provide which general framework of these tools you are referring to, and add them to the appendix. {Editorial

note: Please note these tools with citations but adding them to the appendix isn't necessary.}

Author Response to Reviewer 2 Comment 4: We used a modified version of the Newcastle Ottawa Scale, as has previously been done in systematic reviews of cross-sectional studies. We have clarified this on page 4 and included citation for our approach to risk of bias assessment. Please see our response to [11] and [24].

[26] Reviewer 2 Comment 5: Statistical analyses (Methods): Please describe meta-regression to the readers (as not all readers in CMAJ will be familiar with this form of meta-analysis), and provide a rationale on why meta-regressions are used versus standard meta-analytic procedures for the review. {Editorial note: Only 1-2 sentences are needed.}

Author Response to Reviewer 2 Comment 5: We have made this brief clarification as suggested on page 5: "We present a descriptive summary of associations found for these factors in the literature. Second, we employed meta-regression, a regression-technique of aggregate data which allows for study of the impact of moderator variables on pooled effect size. Specifically, we studied the impact of year burnout data was collected (as a continuous variable), the region (as a categorical variable), and program of residency (medicine vs surgery as categorical variables) on the pooled measure of burnout".

[27] Reviewer 2 Comment 6: Study characteristics (Results): Why did you choose the date range 1987 to 2008? Please provide reason for start date.

Author Response to Reviewer 2 Comment 6: A date range was not chosen. Databases were searched from inception to August 2018. Please see our response to [4].

[28] Reviewer 2 Comment 7: Risk and protective factors (Results): Please provide the categories for relationship status. Marital and relationship status seems to be used interchangeably; please stick to one.

Author Response to Reviewer 2 Comment 7: We have added a definition of this on page 5: "We conducted two secondary analyses. First, we extracted data on reported risk and protective factors, including age, sex, relationship status (single vs. having a partner), depression, level of stress, work hours, frequency of call shifts, job satisfaction, wage/income satisfaction, family/network support, sleep, and level of training". Additionally, we have changed all use to "relationship status" rather than "marital status" as this is more reflective of what is measured in the primary data.

[29] Reviewer 2 Comment 8: Also, you noted that the majority of the studies did not report on associations with burnout for a number of factors; please report reasons why they were not (what are the reasons provided by these authors?)

Author Response to Reviewer 2 Comment 8: Thank you for your insightful suggestion. This was dependent on the study's objective and the research question. For example, if a study sought uniquely to capture level of burnout, the authors did not investigate associated factors. Conversely, some studies had an interest in specific risk factors and therefore these were investigated. While we agree that this data is interesting, it is not possible for us to comment on the study of chosen risk factors for each included study and this is outside the scope of our study. Furthermore, we do not think this additional information would add value or context to the results of this systematic review.

[30] Reviewer 2 Comment 9: Results: Readers may be left wondering why burnout rates are compared in medical and surgical groups only (and not broken down by subspecialty for each larger group).

Author Response to Reviewer 2 Comment 9: We compared burnout rates between medical

vs. surgical training programs for 2 principal reasons: 1) this is often a question of interest<sup>7-9</sup> and 2) we do not have enough data on burnout rates by specialty to study these individually. Any conclusions made by including specific specialty as a predictor in the meta-regression would likely be erroneous. Additionally, we believe that by providing data that burnout rates are not different for surgical vs. medical residents, we may encourage collaboration between training programs to tackle this problem systemically. We have added these points to our discussion on pages 8-9: “Our meta-regression suggests that whether a PMT is a surgical trainee or medical trainee does not significantly explain the heterogeneity in burnout prevalence. While there may exist differences between the two training programs, they are likely small in comparison to other determinants of burnout. The question whether burnout varies between surgical and medical residents has been oft postulated. In fact, it is often hypothesized that surgical residents experience greater stress and harassment<sup>7-10</sup> during their training likely leading to high rates of burnout.<sup>6</sup> Our findings suggest that rates likely do not differ and support alliance of efforts, both policy and research, by Medical and Surgical training programs for a crisis that affects all PMTs equally”.

[31] Reviewer 2 Comment 10: Geographic region (Results): Collapsing countries that have few reports on burnout to increase the sample size for statistical analyses may considerably distort the results, and may provide an inaccurate picture; it also offers little information to readers. Consider providing a narrative review component in the article for these countries, and focus on comparing statistical findings between North America and Europe only. {Editorial note: Adding a few sentences would be helpful.}

Author Response to Reviewer 2 Comment 10: We would like to thank the reviewer for this insightful comment and feedback. Due to the limited data from certain countries, we pooled data from regions/continents of low sample size. As the reviewer suggests, this may potentially provide a slightly inaccurate picture if burnout rates in nearby regions are grossly different. That said, we do not comment about statistical significance between each individual region, but rather comment that we find that region of study overall is a significant predictor of burnout (see Page 7). We have however revised our manuscript in accordance with the reviewer’s suggestion. For the sake of ensuring transparency, we have stated explicitly that “As limited data was available for other regions, stronger conclusions can only be made about burnout among North American and European residents; the prevalence of burnout among European PMTs was 30.8% vs. 51.2% in North America.” Information specific to studies from each of the region is available in Table 1; we prompt the reader to refer to this table on page 7. In order to avoid redundancy especially given the limited word limit, we have not reiterated this in the text.

[32] Reviewer 2 Comment 11: Discussion: Please link back to existing reviews that already reported on the prevalence of burnout among residents; how are your findings different or the same compared to these existing reviews? What does your review add to the existing literature?

Author Response to Reviewer 2 Comment 11: Please see our response to [22].

[33] Reviewer 2 Comment 12: Overall, there is a lot more room for discussion based on the findings of the review. Consider elaborating on different cultural, geographical, and systemic differences among the continents. Even just cultural factors alone may significantly contribute to residents' perception and attitudes of burnout.

Author Response to Reviewer 2 Comment 12: We agree with the reviewer that there are multiple factors that may explain the regional difference in burnout. Please see response to [21]. Certainly cultural, geographical and systemic differences between the continents may play a role in explaining this variation. Given our study is an exploratory study, it is beyond our scope to delineate the specific factors within these regions that is driving the variation. Nonetheless, we do hypothesize some possibilities (e.g., longer paid vacations) but

suggest this as an area for future direction, and we have revised our discussion of this topic to elaborate on the possibility that factors outside the training environment may also be driving this change. For instance, cultural biases may influence one's values and perceptions, thus affecting how an individual responds to the MBI. We have also added the following to page 8: "While it is possible that our findings are biased by methodological considerations such as the fact that MBI may be filled out in a different manner across cultures contributing to the observed variation in prevalence, our findings nonetheless warrant future research to identify cultural and systemic differences that may explain our results, both within and outside the training environment."

[34] Reviewer 2 Comment 13: Interpretation: Page 9, line 50; please provide references on how policies are implemented to focus on wellness.

Author Response to Reviewer 2 Comment 13: Thank you for this comment. We do indeed provide references throughout the text on policies and programs, these include the European time directive (Ref 89), duty hour restrictions overall (Ref 98), among others. We have also added references on program efforts including Ref 27 and 28. We encourage the reviewer to peruse through the citations included in our document.

[35] Reviewer 2 Comment 14: Charts: Page 24, bar chart is missing a caption/title.

Author Response to Reviewer 2 Comment 14: We have added a title and caption to all figures and tables. Please see our response to [16].