Author's response to reviews

Title: Quantitative and Qualitative Perceptions of the 2011 Residency Duty Hour Restrictions: A Multicenter, Multispecialty Cross Sectional Study

Authors:

- William S Tierney (tiernew@ccf.org)
- Rachel L Elkin (elkinr@ccf.org)
- Craig D Nielsen (nielsec@ccf.org)

Version: 4 Date: 23 December 2014

Author's response to reviews: see over
Sunday, December 14, 2014

The Editors of the BMC Medical Education Journal
BioMed Central
Floor 6, 236 Gray’s Inn Road
London
WC1X 8HB
United Kingdom

Dear Editors of BMC Medical Education,

In response to the suggestions and comments of the reviewers for our article “Quantitative and Qualitative Perceptions of the 2011 Residency Duty Hour Restrictions: A Multicenter, Multispecialty Cross Sectional Study,” we have further revised and improved the document. An item-by-item address of their comments follows:

First Reviewer: Jacqueline de Graaf

1. “The authors conclude significant differences in the overall perceptions of DHR across specialty (internal medicine more positive than general surgery) and across position (first year residents more positive than senior residents and faculty). This conclusion is based on 11 GS residents (3 PGY1 + 8 PGY2+) versus 146 IM residents. This may be statistical significance but this conclusion appears not valid because 11 residents cannot be representative for all GS residents. Please respond.”

   • Because we only received responses from GS residents and faculty from one site, for our cross-specialty analyses we opted to focus on a single hospital to keep resident experiences as analogous as possible. As such, for the IM arm of these analyses we only used data from their IM counterparts at the same hospital. Thus, we are comparing 20 GS respondents to 29 IM respondents, rather than the suggested 11 vs. 146. We hope this clarifies the reviewer’s misunderstanding of our data and conclusions for this set of analyses. It is also worth noting that we acknowledge the weakness of having received responses from surgeons in only one of our three sites in our discussion.

2. “The qualitative data are only presented for the whole group together, so no comparison cross specialty and/or cross position and/or cross hospital. Please present who were the responders of the qualitative data. Only IM? The additional value of this study is that qualitative data were gathered and reported. The qualitative data offered new insights into the perceived strengths and weaknesses. However, in the discussion section it is indicated that the qualitative data do not support the quantitative data which is remarkable and interesting. How do the authors explain this? How do you suggest to interpret the data?”

   • We received responses to our open-ended/qualitative questions from IM faculty and residents and well as from GS faculty and residents.

   • In the Discussion section we note that with our qualitative data participants had stronger and more frequent comments about the weaknesses of the DHR as opposed to its strengths, suggesting either that respondents feel more negatively about duty hours or that there is a general culture within graduate medical education/residency programs to emphasize the weaknesses of the DHR. We originally suggest that this may represent a selection bias with regards to which respondents elected to participate in the qualitative portion
of the survey, or that this is reflective of a disconnect between the quantitative and qualitative arms of the survey. To clarify this, we have changed the conclusion of our discussion section to read as follows: “Therefore, a discrepancy exists between the qualitative and quantitative feedback received. This may represent a selection bias with regards to who elected to respond to the qualitative questions or it may indicate a disconnect between quantitative and qualitative perceptions of the DHR. Whether individuals simply find more specific critiques than positive comments when considering the DHR or people with critiques are more likely to complete qualitative questions is not clear from this research. An alternative explanation is that this discrepancy between our qualitative and quantitative data originates because open-ended qualitative questions allowed respondents to comment on some negative aspect of the DHR which was not represented in our quantitative survey items. However, after careful review of all qualitative data we were unable to identify such a factor and this is an unlikely explanation.”

3. Line 66: were interns included too? Where are the data?
   - Interns, or PGY1 residents, were included and their data is included throughout the analysis alongside the PGY2+ residents. We added PGY designations to define interns and other residents to avoid confusion.

4. Line 82: what do the authors mean by ‘residency program statistics were obtained from PD’s directly’?
   - We clarified this by stating that this data was gathered via email correspondence with the Program Directors of each involved residency program. “Residency program statistics” refer to the number of residents in each participating program, as well as the number of core faculty who teach in each residency program.

5. Line 100: resident response rate was 49% (168/341) – did the response rate differ between GS and IM residents? how high was the faculty response rate? Only 11 GS residents are included (of whom only 3 PGY1) - in which hospital did they work?
   - We elected not to report the general surgery response rate because our sample of surgeons was so small. The faculty response rate was not calculated because faculty were contacted via departmental listservs kept confidential from study staff to protect respondents’ anonymity. We added the hospital name (hospital B) into our cross-specialty comparison.

6. Line 141: qualitative data are described for all the whole group, including both residents and staff of GS en IM (and also 12 ‘other’ respondent). Please specify who completed the qualitative data survey. Are there any cross-specialty and / or cross-position and/ or cross-side differences in the qualitative data?
   - As above, we had qualitative responses from residents and faculty in both IM and GS.
   - Our sample size for the qualitative data was too small to be able to make valid cross-position, cross-specialty, of cross-institutional comparisons. As such, our analyses reflect the pooled qualitative data from all respondents.

7. Line 166: you state that individuals in the GS residency program reported more concerns about the DHR than those in IM in the quantitative arm of the study. This is not reported in the result section: here is only referred to the data in Table 5. Unfortunately the cross-specialty comparison could only be performed at 1 hospital – which hospital?
   - This point is indeed mentioned in the Results section; please refer to the second sentence of the “Cross-specialty comparisons” subheading of the Results section. These comparisons were made from the participants affiliated with Hospital B. We have added this information to the results section for clarity.
8. The authors do acknowledge that only 1 GS program participated limiting the power of the cross-specialty comparison. Additionally, no cross-site comparison is possible. Again, the number of responders in the GS program is so low and heterogeneous (PGY1, PGy2+ and faculty) that conclusions seem invalid.

- While we acknowledge that the study is weakened by the relatively low number of GS participants we were still able to reach viable numbers of participants to make comparisons between GS and IM at one site. We included 20 GS respondents including both residents and faculty and stated this in our methods section specifically. Statistically, all comparisons were carried out using appropriate statistical tests for small sample sizes with non-normal distributions. As such, we were limited in our conclusions and only able to compare the entire GS respondent group to the IM respondent group at one site. However, we feel that our results and conclusions reflect this accurately and that a comparison between GS and IM groups adds an important and statistically valid finding to our study.

9. Line 211: ‘felt’ should be deleted

- This word has been deleted.

10. Line 216 - 221: this conclusion is not supported by the data but is the suggestion of the authors.

- This is actually a restatement of the qualitative survey responses. It was unclear in the original document and has been modified to specifically state that these are the opinions of the respondents.

11. Table 1: only n=11 (7%) GS residents are included versus 147 IM residents. How about the n=10 ‘other residents’. I suppose they are from different specialties? Did you exclude these ‘10 others’ from the analysis?

- These 10 were residents from other specialties who were rotating on GS or IM teams when the survey was distributed. They were excluded from the analysis.

12. Residents who trained at osteopathic medical schools reported significantly more negative views of DHR than those who had trained at allopathic or international medical schools, suggesting an influence of undergraduate medical training. The authors put this forward throughout the manuscript as one of the main findings. Is it possible that this is a chance finding? Do the authors suggest that this result implicates that undergraduate medical education should be aware of and discuss RDH when guiding students in career development? Was there any cross-specialty difference in what medical school they visited? Please include a short explanation what is meant by osteopathic versus allopathic medical school because this is not directly obvious for readers outside the USA/Canada.

- This was the one completely unexpected finding of our study. With small sample sizes one has to be cautious, but using appropriate non-parametric approaches (and with the guidance of a PhD-level biostatistician) we found statistically valid differences based upon type of undergraduate medical education even with our most conservative testing methods. We hesitate to suggest anything definitive based on this finding because it is both novel and without a simple explanation. However, this finding is important to add to the literature as it may increase interest in future research of the impact of undergraduate medical education on residency satisfaction.

- A brief explanation on the differences between allopathic and osteopathic training has been added to the manuscript.

13. In the discussion I miss discussion about the sleeping hours and how DHR affect career development. About both items the authors do report data in the result section.
• We elected not to comment on our sleep data because it was uniform across all groups and we felt that these findings merited no additional discussion in the discussion section. Additionally, other studies with much larger sample sizes report very similar data; our study simply supports their findings.

14. Line 198: It is unclear to me why and how the length of content analysis is associated with stronger opinions of the weaknesses than the strengths. Is there any evidence for this reasoning? I think it is an association with no direct ‘causality’.

• We modified this line to state that participants had “stronger and more plentiful comments on the weaknesses of the DHR than the strengths.” This is simply an observation that people wrote a lot more content for their critique of the DHR than in praise of it. While this is not a direct assessment tool we believe that it does inform the reader of the sort of responses we received in our surveys: ones which placed more emphasis on the negative aspects of the newest DHR in form, length, and number.

**Second Reviewer: Brian C. Drolet**

1. When and why was this study conducted? There have been larger studies of the exact same nature published. This is well written, and conducted well (reasonable response rate, data analysis, etc).

• Data were collected in 2012 during the first year of the new DHR and data analysis was completed in early 2013. The study’s publication was delayed because several larger studies conducted at similar times came out while our manuscript was under review. Our paper thus was deferred by two journals based on this issue. We submitted our manuscript to BMC Medical Education on January 7, 2014.

• Our study was conducted with the intention to examine the impact of hospital setting (institution), specialty, and position/level of training on perceptions of the newest DHR. We also wanted to incorporate a qualitative data arm to provide richer insights into an issue that usually is addressed with purely quantitative approaches.

2. The comment that this is the first study to include qualitative data imposes a higher standard for the discussion of methodology; this sounds like content analysis to me, but 2 reviewers assigning codes, as described is not very descriptive. Does this component add anything to the general body of knowledge regarding duty hours?

• You are correct that this is content analysis, more specifically conceptual analysis. We have made this clearer in the manuscript.

• Although some of the themes from our qualitative data echo those reported in previous studies, we do believe that there are lessons to be learned from this arm of our study. First, some of the strengths and weaknesses derived from the data, including increased independent study time and adoption of a shift mentality, are not reported elsewhere to our knowledge. We also received several actionable and specific suggestions addressing each of the three primary goals of this most recent ACGME DHR reform: education, patient care, and quality of life. While some ideas apply to DHR reform more generally, others were more program-specific. More and more data are emerging that the “one size fits all approach” is insufficient to meet the needs of every training program. Program directors will need informed input to be able to adapt their programs to best meet the needs of their trainees and teaching staff and to address each of the three arms of DHR reform. Our data suggest that residents and teaching faculty and ready and well-equipped to provide such input, which could be captured through a short survey, focus groups, or less formal discussions.

3. Overall, my review is limited and my recommendation is negative because I do not see any additional value to the publication of this manuscript. Should the editors find otherwise, I would not recommend significant changes to the manuscript other than some further description of the quantitative data review process.
• We appreciate the thoughtful feedback from both reviewers and regret that this study was not being published more expediently. However, we believe that its impact within the graduate medical education community is still meaningful and relevant, as we discuss in our responses above. We hope that Dr. Drolet and the editors of BMC Medical Education agree with this belief after further review of our revised manuscript.

Sincerely,

William S. Tierney
Corresponding Author